DOCUMENT RESUME

ED 470 298 TM 034 520

AUTHOR	Smith, Julia B.; Lee, Valerie E.; Newmann, Fred M.
TITLE	Instruction and Achievement in Chicago Elementary Schools. Improving Chicago's Schools.
INSTITUTION	Consortium on Chicago School Research, IL.
PUB DATE	2001-01-00
NOTE	51p.
PUB TYPE	Reports - Research (143)
EDRS PRICE	EDRS Price MF01/PC03 Plus Postage.
DESCRIPTORS	Academic Achievement; *Elementary Education; Elementary School Students; *Interaction; *Learning; *Standardized Tests; *Teaching Methods
IDENTIFIERS	*Chicago Public Schools IL

ABSTRACT

This study focused on the link between different forms of instruction and learning in Chicago, Illinois, elementary schools. It used teachers' survey reports about their instruction in the 1997-1997 school year and linked these reports with achievement gains. The study tested the common assumption that the nature of standardized assessments requires that teachers who want to enhance their students' test scores should make extensive use of the classroom drill and practice activities associated with didactic instruction and review rather than more interactive teaching. Student information was available for 110,775 students. The composite sample for teacher surveys was 5,586 from 384 schools. The study found strong empirical evidence that in Chicago's elementary schools the instructional approach teachers use influences how much students learn in reading and mathematics. Interactive teaching methods were associated with more learning in both subjects. Findings call into question the assumption that low-achieving economically disadvantaged students are best served by teaching that emphasizes didactic methods and review. The study also found important relationships between teachers' professional preparation and the presence of key organizational supports within their schools. These findings support policy efforts to education teachers on how to educate teachers to use interactive methods with all their students, to provide opportunities for teachers to engage in dialogues with their colleagues, and to encourage principals to provide strong instructional leadership. An appendix contains details on data sources and samples used. (Contains 14 figures, 16 endnotes, and 21 references.) (SLD)

Instruction and Achievement in Chicago Elementary Schools

Julia B. Smith

Valerie E. Lee

Fred M. Newmann

January 2001

Improving Chicago's Schools

Consortium on Chicago School Research

1313 East 60th Street
Chicago, Illinois 60637
773-702-3364

Executive Summary

Although there are many ways to improve student learning, improving instruction is surely near or at the top of any list of educational reforms. This study focuses on the link between the different forms of instruction and learning in Chicago elementary schools. It makes use of teachers' survey reports about their instruction in the 1996-97 school year and links these reports with student achievement gains. The study tests a common assumption: That the nature of standardized assessments requires that teachers who want to enhance their students' test scores should make extensive use of the classroom drill and practice activities associated with didactic instruction and review rather than more interactive teaching.

This study provides strong empirical support that "instruction matters." We found clear and consistent evidence that in Chicago's elementary schools the instructional approach teachers use influences how much students learn in reading and mathematics. Moreover, interactive teaching methods were associated with more learning in both subjects. Our findings call into serious question the assumption that low-achieving, economically disadvantaged students are best served by teaching that emphasizes didactic methods and review. We also found important relationships between teachers' professional preparation and the presence of key organizational supports within their schools, and their use of the more effective interactive methods. These findings support policy efforts to educate teachers on how to use interactive methods with all their students, to provide opportunities for teachers to engage in dialogues about instructional practices with colleagues in their schools, and to encourage principals to provide strong instructional leadership.

We conclude that efforts to engage all students in deeper and broader thinking about subject matter are a hallmark of "good teaching," and that Chicago students' achievement could improve further if teachers across the school system were encouraged to achieve a better balance among their use of review, interactive teaching, and didactic teaching practices.

Instruction and Achievement in Chicago Elementary Schools

Julia B. Smith
Oakland University

Valerie E. Lee
University of Michigan

Fred M. Newmann
University of Wisconsin

January 2001

John Booz

Foreword

In 1993, Ambassador Walter Annenberg announced a $500 million challenge grant to improve public education in the United States. Cities wishing to receive a portion of that grant were invited to submit proposals describing how the funds would be used to stimulate educational innovation and collaboration in their public school systems. A group of Chicago school reform activists and education stakeholders, including parents, teachers, principals, community leaders, and foundation officers, organized to write a proposal to include Chicago among the sites receiving a grant. They were successful. In January 1995, the Annenberg Foundation awarded a five-year grant of $49.2 million to establish the Chicago Annenberg Challenge. An additional $100 million in matching funds was pledged by local donors.

The Chicago Annenberg Challenge was organized to distribute and manage these monies among networks of schools and external partners throughout the city. Its mission is to improve student learning by supporting intensive efforts to reconnect schools to their communities, restructure education, and improve classroom teaching. The Chicago Challenge funds networks and external partners that seek to develop successful, community-based schools that address three critical education issues through whole-school change: school and teacher isolation, school size and personalism, and time for learning and improvement. More than half of Chicago's public schools will have participated in an Annenberg-supported improvement effort by the end of the grant period in 2001.

This report is part of a series of special topic reports developed by the Chicago Annenberg Research Project. This series focuses on key issues and problems of relevance to the Chicago Annenberg Challenge and to the improvement of Chicago public schools generally. It complements a series of technical reports that focus specifically on the work and accomplishments of the Chicago Annenberg Challenge. Among the topics examined to date in the special topics report series are

the quality of intellectual work in Chicago elementary schools; social support, academic press, and their relationship to student achievement; and this report, *Instruction and Achievement in Chicago Elementary Schools.*

The work of the Chicago Annenberg Research Project is intended to provide feedback and useful information to the Chicago Challenge and the schools and external partners who participate in its efforts to improve educational opportunities for Chicago's children and youth. This work is also intended to expand public discussion about the conditions of education in the Chicago Public Schools and the kinds of efforts needed to advance meaningful improvements. This effort to stimulate new avenues of discussion about urban school improvement is an important aspect of Ambassador Annenberg's challenge to engage the public more fully in school reform.

Acknowledgments

Many organizations and individuals contributed to this report. The authors are deeply grateful for their assistance. The Chicago Annenberg Challenge provided the primary funds for the study and for the development and distribution of this report.

We would like to thank staff in the Office of Accountability at the Chicago Public Schools for their ongoing assistance and especially for providing the test score data essential for conducting this study. In particular, Joseph Hahn, Andrea Ross, Gudelia Lopez, and Sandra Storey helped us with this data. We are also very grateful to the computing staff at the Consortium, who were extremely helpful in bringing specificity to the measures. Consortium Director Anthony S. Bryk provided essential guidance through the preparation of this study.

Other members of the Lead Team of the Chicago Annenberg Research Project gave important counsel during all phases of the report, from planning through publication; special thanks go to Mark Smylie, Stacy Wenzel, Karin Sconzert, and Rodney Harris.

Sandra Jennings, Sarah-Kay McDonald, and Carolyn Saper provided design, production, and editorial services. The photographs are by John Booz.

John Booz

I. Examining Instruction

It is 10:00 a.m. in a Chicago elementary school, time for language arts. In Classroom A, a third-grade teacher pronounces the vocabulary words for the day carefully and provides definitions for each word as she writes them on the board. After the students copy down the words and definitions, they turn to worksheets to practice using the new words in sentences. When the students have finished their worksheets, the teacher asks everyone to open their reading books to the same page, and the students take turns reading passages aloud. Toward the end of the lesson, the teacher draws students' attention to places in the story where the words are used, and asks yes/no questions to check comprehension.

In Classroom B, the teacher distributes lists of the same vocabulary words to her third graders. She divides the students into groups to undertake projects related to the words. One group uses the dictionary to locate definitions. Another group writes a story using the words. A third group develops a crossword puzzle, working up clues that will use the words as answers. When the groups complete these activities, they trade their results with one another, checking other students' work and critiquing the product. During the language arts period, the teacher moves from group to group, coordinating activities and providing assistance when students ask for help. Over the entire period, each group will be involved in each activity.[1]

The content for these lessons is identical in both third-grade classrooms. That day's lesson plans for both teachers indicate that their classes are at exactly the same place in the year's progression of topics. The ways these lessons are taught in these two classrooms, however, differ dramatically. This comparison illustrates what has become a critical concern for school reform: how teachers should organize their lessons so that students will learn the knowledge and skills that the teachers wish to impart. The issue, in short, is how best to approach instruction to maximize student learning.

Different Approaches to Instruction

There has been abundant discussion among policy makers, educational researchers, practitioners, parents, and the interested public about what constitutes good teaching. In discussing the criteria for these judgments, people frequently split into two camps. One camp focuses on students' knowledge of content, which usually involves teaching through lectures, drill, practice, and worksheets. Instruction of this type is typically delivered through tightly structured exercises in which the whole class engages the content at the same time and in the same way. The exercises require students to memorize, recite, and demonstrate facts, definitions, and procedures.

Members of the other camp emphasize the equal importance of process and the advantage of trying to engage students in the lesson, sometimes by relating its content to their lives. To deliver instruction of this second type, teachers typically spend less time on whole-group activity and more time coaching students (individually, in small groups, or sometimes the whole class) through discussion, hands-on activities, and projects that students themselves have a role in defining.

Different labels have been applied to instruction by the two camps. The first has sometimes been called traditional, teacher-centered, directive, highly structured, or didactic instruction. The second type has been labeled progressive, student-centered, constructivist, authentic, or interactive instruction. Although precise definitions of instruction that carry these labels differ, in this report we use the terms *didactic* and *interactive* to distinguish between the first and the second approach to instruction. The two approaches differ in many ways, including the behaviors in which teachers and students engage and the intellectual demands made on students. Because writers and practitioners take different perspectives on the definition of the most effective approaches to instruction, the language used to describe these two camps is neither widely accepted nor very precise.

Study Advances Prior Research on Effective Instruction

Prior research has documented substantial achievement benefits, and no consistent disadvantages, for students exposed to the kinds of teaching characterized in this report as interactive instruction. These studies include research on the teaching of mathematics, reading, and writing to disadvantaged students (Knapp, Shields, and Turnbull 1992); teaching mathematics in grades one, two, and eight (Carpenter, Fennema, Peterson, Chiang, and Loef 1989; Cobb, Wood, Yackel, Nicholls, Wheatley, Trigatti, and Perlwitz 1991; Silver and Lane 1995); teaching reading in grades one, two, and eight (Tharp 1982), teaching mathematics and science in high school (Lee, Smith, and Croninger 1997); and teaching social studies in high school (Levin, Newmann, and Oliver 1969). Many of these are relatively small-scale studies that have applied diverse definitions for didactic and interactive instruction, and studied different grade levels and different subjects.

This study represents the largest base of data from urban schools and students that has been brought to bear on this concern to date. The study includes test scores from more than 100,000 students in grades two through eight and surveys from more than 5,000 teachers in 384 Chicago elementary schools. Both the size of the samples and the depth of the data set should lead to more generalizable conclusions than prior studies have reported.

Nevertheless, the issues raised by the two sides have generated heated controversy. Discussions of these two approaches often suggest that they are mutually exclusive. Moreover, both professional educators and policy advocates often argue in favor of one approach and disparage the other. However, in practice good teachers probably combine both approaches, depending on the content to be taught and the abilities and interests of their students. Table 1 attempts to characterize the behaviors of teachers and students under each approach. Of course, other authors might offer somewhat different emphases and distinctions.[2]

Table 1

Contrasting Didactic and Interactive Instruction Components

Didactic	Interactive
Teachers usually . . .	***Teachers usually . . .***
· Lecture or demonstrate to students	· Coach, listen, and guide students
· Pose questions that ask for single, short answers	· Pose questions that ask for explanations and which may have multiple answers
· Assess students on correctness of answers	· Assess how students arrived at answers
· Determine what students will study	· Provide choices in what students study
Students usually . . .	***Students usually . . .***
· Listen to teacher and recite answers	· Discuss answers and ideas with teacher and peers
· Try to repeat the knowledge they have been taught as it was transmitted	· Try to apply, interpret, and integrate knowledge into prior understanding
· Rarely choose what questions or topics to study	· Frequently choose what questions or topics to study

Didactic instruction. When using didactic instruction, the teacher functions as the major source of knowledge—the expert. Because knowledge is usually seen as conclusive and objective, lecturing is often considered the most efficient method of teaching. The teacher typically delivers information this way to the entire class, possibly embellishing lectures by performing demonstrations. To assess students' understanding of the content, the teacher often poses questions that require short answers. These questions usually have a single correct answer that the teacher knows in advance, because the knowledge to be conveyed is considered authoritative and not subject to debate. Students' understanding is assessed on the correctness of answers to questions that are based on knowledge that has been previously communicated. Assessment of student learning in a didactic instructional format occurs either in oral recitation exercises, or, more often, in writing on short-answer or multiple-choice tests.

In didactic instruction, students' time in class is typically spent: (1) listening to the teacher, (2) reciting answers to questions, or (3) practicing skills and information retrieval by completing worksheets or exercises. An example of this activity can be observed during a lecture period. Upper-elementary school students may be taking notes on what the teacher is saying, writing down what is on the board, responding to questions put to them by the teacher, or practicing what they have just heard by completing written exercises supplied by the teacher. Rarely are students asked or allowed to choose the questions or topics to study; rather, the topics and questions are supplied by the teacher. All students generally engage in the same tasks at the same time, and those tasks focus predominantly on replicating skills and knowledge presented by the teacher or drawn from a textbook. Assessment is based on the expectation that students can repeat the knowledge they have been taught, in the same or similar form in which it was transmitted.

Interactive instruction. In interactive instruction, the teacher's role is primarily one of guide or coach. Teachers using this form of instruction create situations in which students encounter knowledge in ways that provoke them to ask questions, develop strategies for solving problems, and communicate with one another. Teachers pose questions to students either individually, in small groups, or to the entire class. The questions may not always be determined prior to the lesson, and these problems may have more than one answer that could be considered reasonable or correct. As a result, students are often expected to explain their answers and discuss how they arrived at their conclusions. Teachers who use interactive instruction encourage their students to relate class work to their own experience. These teachers usually assess students' mastery of knowledge through discussions, projects, or tests that demand explanation and extended writing. Besides content mastery, the process of developing the answer is also viewed as important in assessing the quality of the students' work.

In interactive instruction, the teacher's role is primarily one of guide or coach. Teachers using this form of instruction create situations in which students encounter knowledge in ways that provoke them to ask questions, develop strategies for solving problems, and communicate with one another.

In classrooms that emphasize interactive instruction, students discuss ideas and answers by talking, and sometimes arguing, with each other and with the teacher. Students work on applications or interpretations of the material to develop new or deeper understandings of a given topic. Such assignments may take several days to complete. Students in interactive classrooms are often encouraged to choose the questions or topics they wish to study within an instructional unit designed by the teacher. Different students may be working on different tasks during the same class period.

Combining the approaches. The didactic and interactive approaches reflect important contrasts in teachers' and students' behaviors. However, probably very few teachers use only one approach. The comparisons between the two approaches made in Table 1 may exaggerate or over-simplify differences in teachers' actual practices. Moreover, virtually all teachers spend time reviewing previously taught skills and content. Therefore, it seems reasonable that many teachers combine practices from each approach, depending upon the demands of the curriculum and the types of students. Although the approaches may be combined, it is helpful, nevertheless, to learn more about the effectiveness of each approach in its distinctive form. For example, some researchers and policy advocates worry that some forms of interactive instruction may be harmful to economically disadvantaged students who have fewer opportunities to learn basic skills in their homes.[3]

Review. Regardless of a teacher's emphasis on didactic or interactive practices, all teachers face the issue of how much to review previous lessons before moving on to new material or to deepen students' understanding. There are several reasons why teachers may feel that review is necessary. Because of mobility (i.e., new students joining the class), absence, or tardiness, some students may not have been exposed to important content. In addition, students' in-class responses may indicate that many class members have not yet mastered content that was taught in previous classes or should have been learned in previous grades. Moreover, even if the teacher were confident that prior lessons were mastered, students may have forgotten important knowledge and skills because they had little opportunity to use them.

Although repeating information may help to reinforce learning, invariably such review takes time away

from developing more complex understanding and introducing new material. Furthermore, students who do not need the review may become bored and less engaged in learning. Thus, spending much time on review may decrease, rather than increase, how much the student body as a whole learns. Studies of instruction seldom explore how the amount of review influences student achievement. Moreover, the current press for students to score well on standardized tests may increase teachers' perceptions of the need for review. This study of how instruction influences achievement has, therefore, included an examination of review.

Research Focus

To clarify the debate about which approach best serves diverse student needs, we turned to the Consortium's teacher surveys and student achievement data to investigate instruction in the city of Chicago. In these analyses, we focused our inquiry around three interrelated questions. First, we wanted to know which schools and classrooms, and particularly which students, receive more or less didactic, interactive, and review-oriented instruction. Our concern here, as in past Consortium reports, is to explore which students experience which types of instruction.

Once the patterns of exposure are clear, we then explore the impact of these different kinds of instruction on student learning. We are interested in how instructional approaches affect students' scores on standardized assessments, such as the Iowa Test of Basic Skills. A frequently encountered belief is that the highly structured (i.e., true/false and multiple choice) nature of these tests necessitates the extensive classroom drill and practice found in the didactic approach. This investigation tests that assumption, taking into account both student and school characteristics which might otherwise bias observed relationships.

Finally, we examine teacher and school characteristics that might be used to promote more successful instruction. Here we focus on issues of teacher training, professional development, and school professional organization which might link to uses of different types of instruction.

How We Did This Study

Data for this study come from two sources: surveys and test scores. Information about students' exposure to different forms of instruction, and about the characteristics of teachers, classroom conditions, and schools were drawn from surveys of Chicago's elementary school teachers conducted in spring 1997. In these surveys, teachers described the frequency of their use of several instructional practices during school year 1996-1997, as well as reporting on conditions in their classrooms and schools. For many years, concerns have been raised about the validity of survey self-reports. Studies have shown that respondents tend to over-report practices and attitudes considered socially desirable and under-report those considered more negative. However, in the 1997 Consortium survey used in this study, teachers' responses revealed considerable variance. They also did not cluster toward interactive methods, which now tend to be more widely endorsed in the rhetoric of school reform.

The second data source was students' annual achievement scores on the Iowa Test of Basic Skills (ITBS) in reading and mathematics. Every Chicago student completes these standardized tests during the spring of each year. The measures of achievement for this study were students' scores on the ITBS in 1996 and 1997. Achievement gain, or learning, was computed as the difference between the scores in each subject over that school year.

The major analytic method used for the analyses in this report is Hierarchical Linear Models (HLM—Bryk and Raudenbush 1992). HLM is the most appropriate statistical technique for analyzing data and research questions with a multilevel structure. In this study, both students and teachers were nested in grade levels and in schools. Because Consortium surveys do not include identifications of individual classrooms, we could not link students directly to their classroom teachers' reports about the use of instructional approaches. Rather, these data are linked at the grade level (i.e., we know both the students' and the teachers' grade level within a school).

Thus, we estimated students' exposure to the three instructional approaches taking into account characteristics of the students themselves, their classrooms, and their schools. Specifically, we included measures of students' gender, minority status, poverty status, and retention history; classroom grade levels, ability level, problem behaviors, and absenteeism; and schools' average income, average achievement in 1996, racial or ethnic composition, school instability, and size in our HLM models that estimate instructional effects on achievement. As a result, all effect estimates presented here are "net of" or "controlled for" these other factors. Further technical details about the study, including samples of students, teachers, and schools used in the analyses are provided in the appendix. The appendix also provides information about the measures, including how they were constructed, the actual survey items included in them, and the structure of the HLM models.

II. The Equity Question: Who Gets Exposed to What Kinds of Instruction?

Many factors influence teachers' instructional decisions, some of which are associated with conditions in the classrooms and schools. A teacher decides from day to day (and sometimes from moment to moment) how best to present new content and review old content, how tasks should be structured, how to coordinate questions from students, and how to negotiate the transitions from topic to topic. As a result, teachers may use (and students may experience) very different kinds of instruction. In this section we explore how instructional usage is linked to classroom and school contexts.

The Classroom Context

Among the characteristics of elementary school classrooms related to instruction, this report considers four: the classroom grade level; the achievement level of the class compared to the school average; the prevalence of behavior problems; and the amount of disturbance caused by irregular attendance. Our aim here is to understand the classroom conditions that might influence the pedagogic approaches that teachers use.

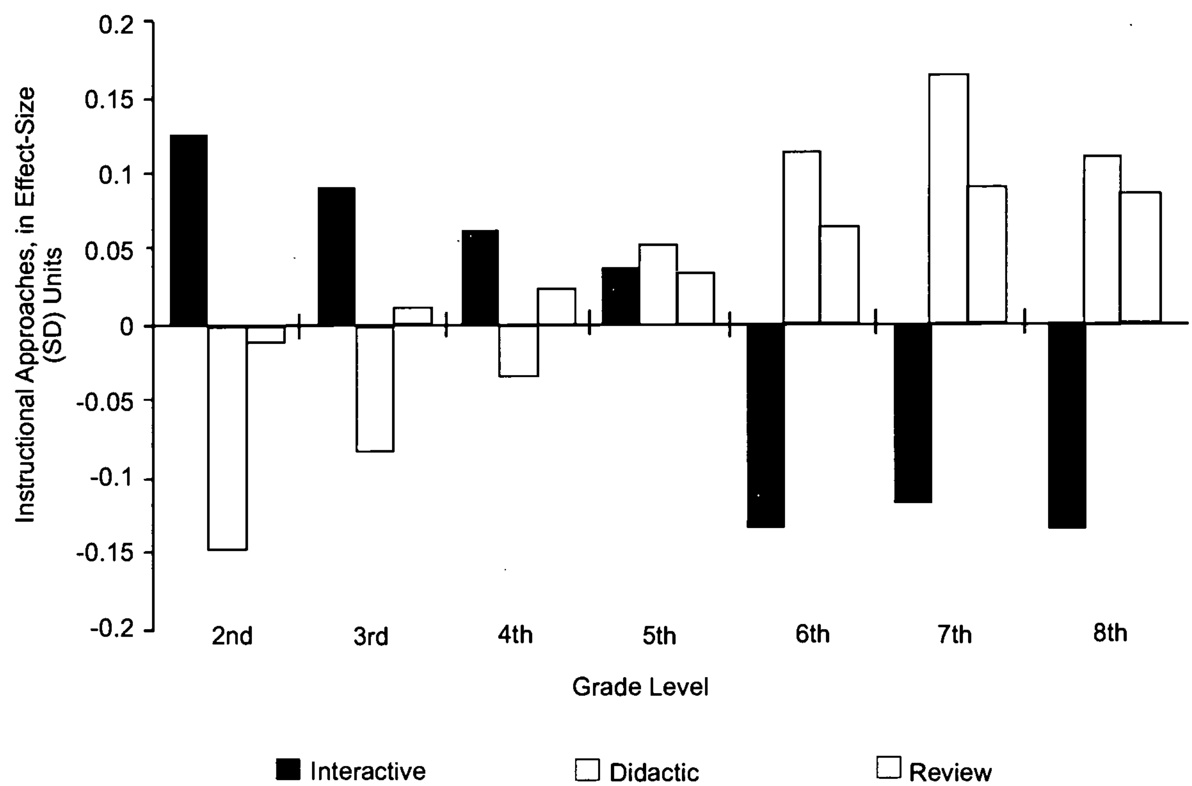
We focus on the frequency of interactive, didactic, and review instruction under these four classroom conditions. All of our analyses of associations between the three forms of instruction and the four classroom conditions also take into account or "control for" the other characteristics of teachers and classroom contextual factors considered here. Because the three types of instruction were measured with composite variables, we present our results in an effect-size metric, so that readers may make comparisons between the size of effects across the several measures.

What Are Effect Sizes?

The results of several analyses in this report are presented in a metric called "effect sizes." For Figures 1-4 and 12-14, results are in this metric. Effect sizes are actually the same as standard deviation (SD) units. There are at least two advantages in using effect sizes to present research results. First, they represent a standard metric that allows those who read research to make comparisons across analyses that focus on outcomes that may actually be measured in different metrics. Second, because the effect-size metric is more and more commonly used to present research results, it is appreciated by readers because the magnitudes of effects may be interpreted substantively, above and beyond their level of statistical significance. In general, researchers agree that effect sizes of .5 SD or more are large, those between .3 SD and .5 SD are medium, those between .1 SD and .3 SD are small, and those below .1 SD are trivial (Rosenthal and Rosnow 1984, p. 360). The effect sizes shown in this report range from very small to quite large, although all that we report are statistically significant.

Figure 1

Didactic Instruction Increases and Interactive Methods Decline Across Grade Levels



Grade level. Figure 1 shows that the frequency with which teachers use instruction described as interactive, didactic, and review varied considerably and consistently by the grade level of the class. Most striking is a comparison between interactive and didactic instruction. Teachers used interactive instruction most frequently in the lower grades (grades two through four), but its use declined considerably in the upper grades (grades six through eight). The opposite trend is evident for didactic instruction. Teachers reported using didactic instruction less frequently in the lower grades, more often in the upper grades.[4]

Class achievement level. This aspect was measured by teachers' reports of whether their students were above, at, or below grade level. Figure 2 suggests that teachers used interactive instruction more frequently in classes that they described as at or above grade level, whereas they used didactic instruction least often in such classes. It is understandable that the largest differences in instructional usage for classes of different achievement levels occurred in the category of review. Teachers used review most frequently in below-grade-level classes, but very infrequently in classes

Figure 2

Interactive Instruction is More Common in Classrooms with High Levels of Prior Achievement; Didactic and Review Practices are More Common in Low-Level Classes

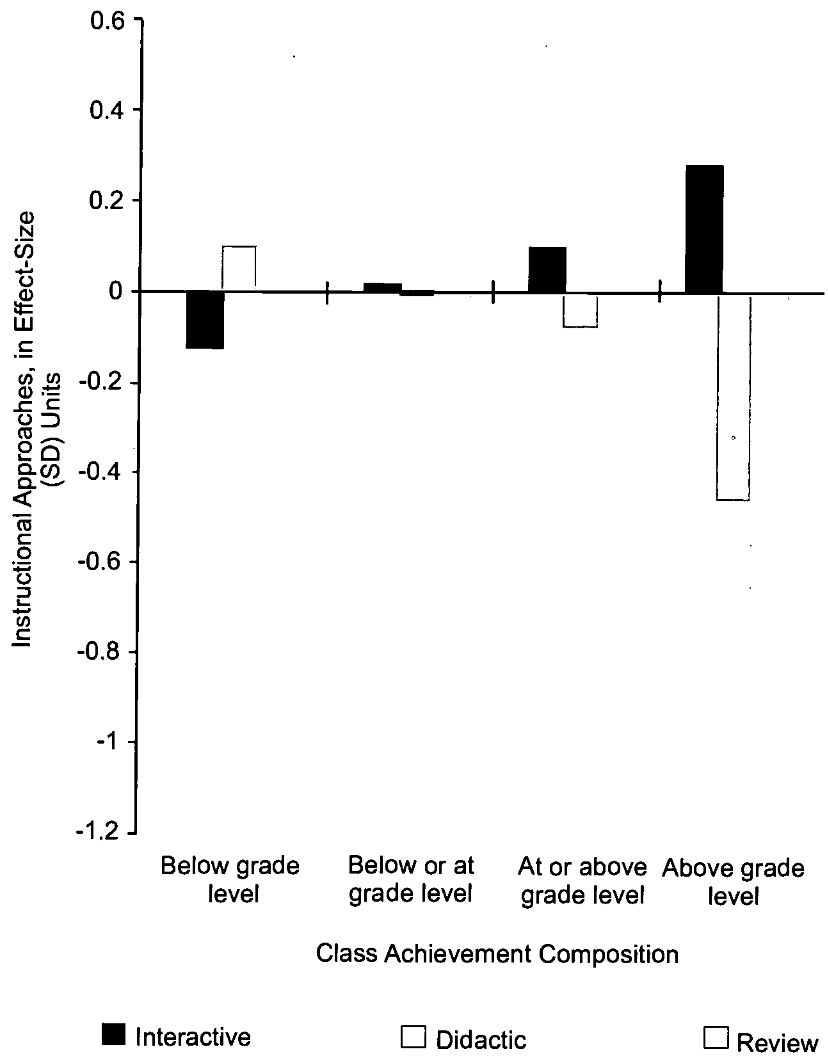
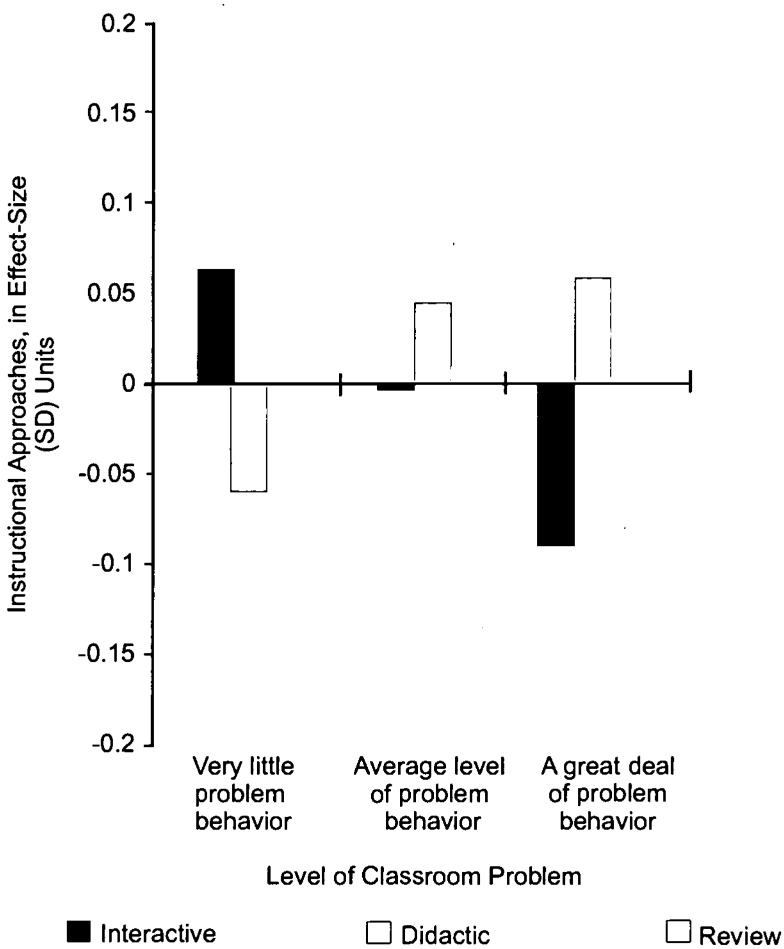


Figure 3

A Greater Emphasis on Review Occurs in Problem Classrooms



where students were above grade level. Classrooms composed of students who were the furthest behind were most frequently taught with didactic instruction and review.

Problem behaviors in classrooms. Figure 3 focuses on teachers' reports of problem behaviors in their target classes, and how the level of classroom problem behaviors affects the use of instructional approaches. Although differences here were less striking than for grade level and classroom prior achievement level, a trend is nevertheless evident. Didactic instruction and review were more often used in classrooms where problem behaviors were common (the right-hand set of bars in Figure 3); interactive instruction was most commonly used in classroom settings typified by very little problem behavior.

Two interpretations of these results are possible. It could be argued that interactive methods can only be used with "well behaved" children. As a result, teachers eschew more innovative approaches when confronted with difficult classroom conditions. Alternatively, many reformers would argue that these are precisely the kinds of students most in need of interactive teaching. According to this view, classroom problems exist precisely because of dull and dreary instruction. Although our analyses indicate a clear link be-

tween teaching methods and behavior problems, sorting out the causal direction is more complicated.

Attendance problems. Classroom attendance problems are also associated with instructional usage, as Figure 4 indicates. We grouped teachers' reports about attendance into three categories: (1) classrooms where attendance was regular (left-hand set of bars in Figure 4), (2) classrooms where attendance was typical, but somewhat irregular (middle bars), and (3) classrooms where irregular attendance was quite common (right-hand set of bars).[5] In classrooms with irregular attendance, teachers tend to use more didactic instruction and review, and less interactive methods. As discussed above, the same two interpretations of this relationship can be offered here.

The School Context

We now consider whether structural characteristics of schools influence the organization of instruction. For this purpose, we divided school averages of their teachers' reports about the three types of instruction into three categories (high, medium, and low).[6] To show the relationship between instructional usage and school demographic and structural contexts, we display the percent of schools in the "high use" category for each instructional approach.

Figure 4

A Greater Emphasis on Review and Didactic Instruction Occurs in Classrooms with Irregular Attendance

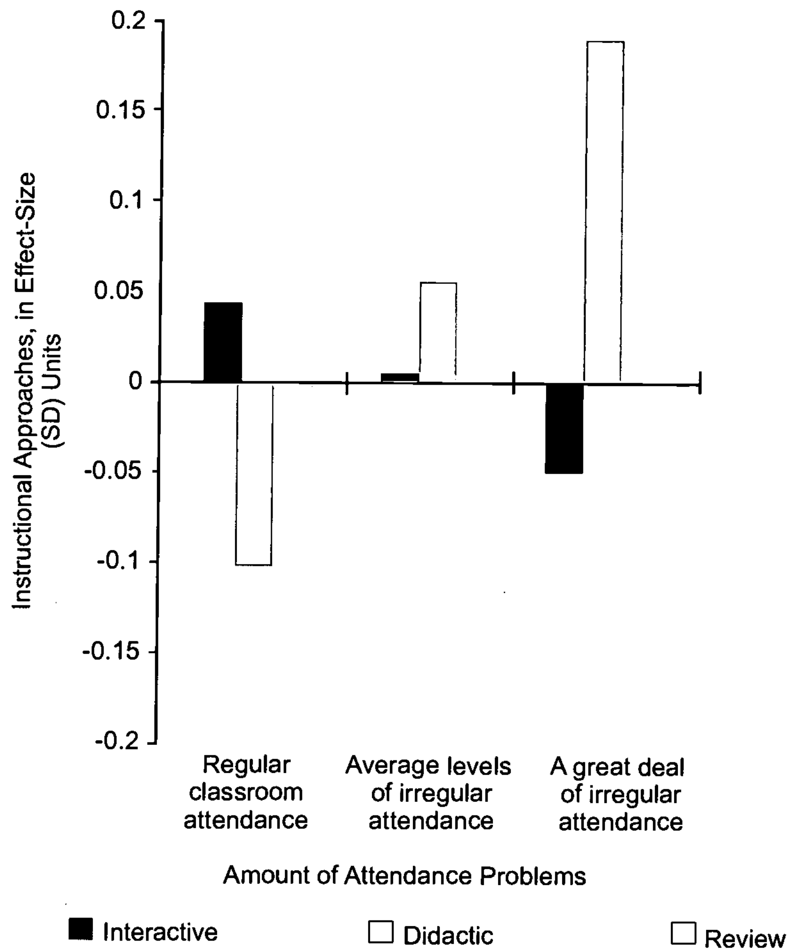


Figure 5

More Interactive Instruction Occurs in Small Schools

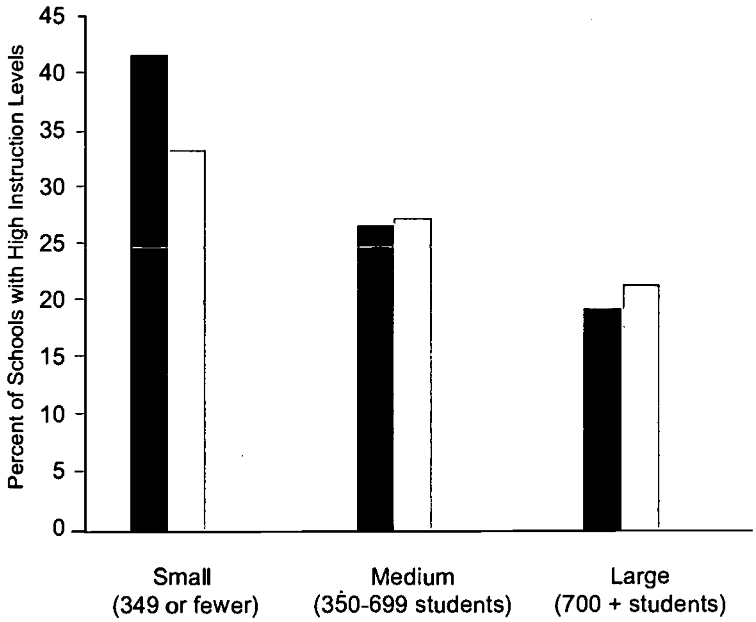
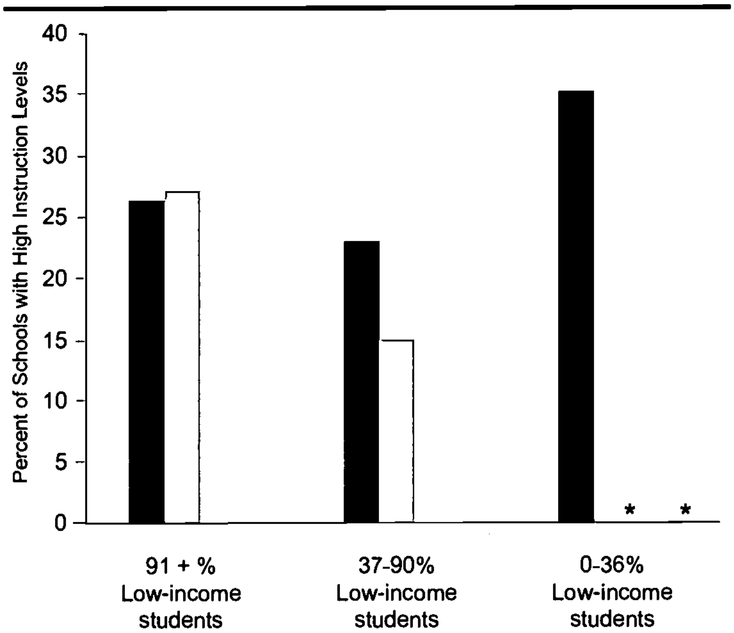


Figure 6

Interactive Instruction More Common in Schools With Fewer Low-Income Students



* No schools in this group reported high use of didactic instruction or review.

School size. Figure 5 indicates that instructional usage, as captured by the percent of schools where each type of instruction is used at a high level, is related to the number of students that the school enrolls. Teachers in smaller schools (fewer than 350 students) report using interactive methods more often and review the least, with didactic usage midway between the two. The exact opposite trend is evident, however, in schools with 700 or more students. Review is particularly common here. In general, as schools get larger, both didactic and interactive instruction decline, whereas the use of review increases.

Enrollment of low-income students. Instructional usage is also linked to the proportion of low-income students a school enrolls, as shown in Figure 6. In schools where almost all students are from low-income families (left-hand set of bars), review is most common. However, for schools enrolling few low-income students (right-hand bars), the use of interactive instruction is high. Interestingly, not a single school in the latter sub-group made extensive use of review or didactic methods.

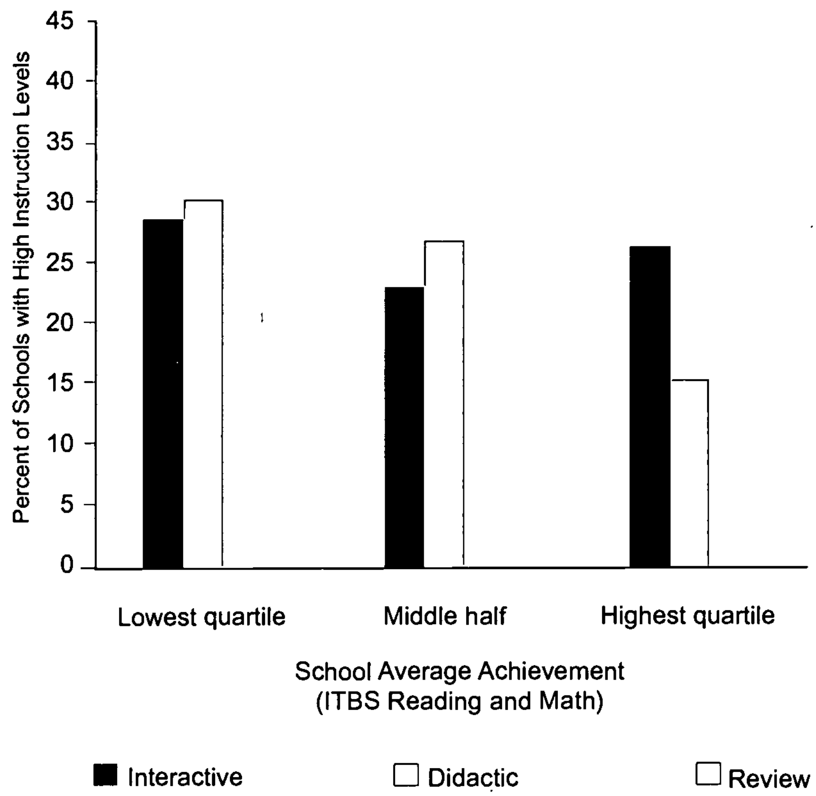
School achievement level. Not surprisingly, results describing instructional usage by school type reflect the findings for classroom achievement shown earlier in Figure 2.[7] As was the case for classroom ability level, the differentiation of

instructional usage was stronger for didactic instruction and review than for interactive instruction (see Figure 7). In general, didactic instruction and review are most common in lower-achieving schools (and classrooms) and less common in higher-achieving schools. However, the use of interactive instruction was unrelated to school achievement level (and was only modestly related to classroom achievement levels).

School racial or ethnic composition. Figure 8 displays the percent of schools using the three types of instruction at a high level for schools with five different racial or ethnic compositions. Contrary to other demographic and structural school characteristics, which showed consistent patterns with instructional usage, few patterns were discernable here. The only major exception is for predominantly African-American schools, whose teachers reported using didactic instruction more frequently than teachers in schools with other racial or ethnic compositions. Interactive instruction was also used more often, and review less, in both racially mixed and integrated schools.

Figure 7

Didactic Instruction and Review is More Common in Schools with Low Prior Achievement Levels



Summary of Contextual Differences in Instructional Usage

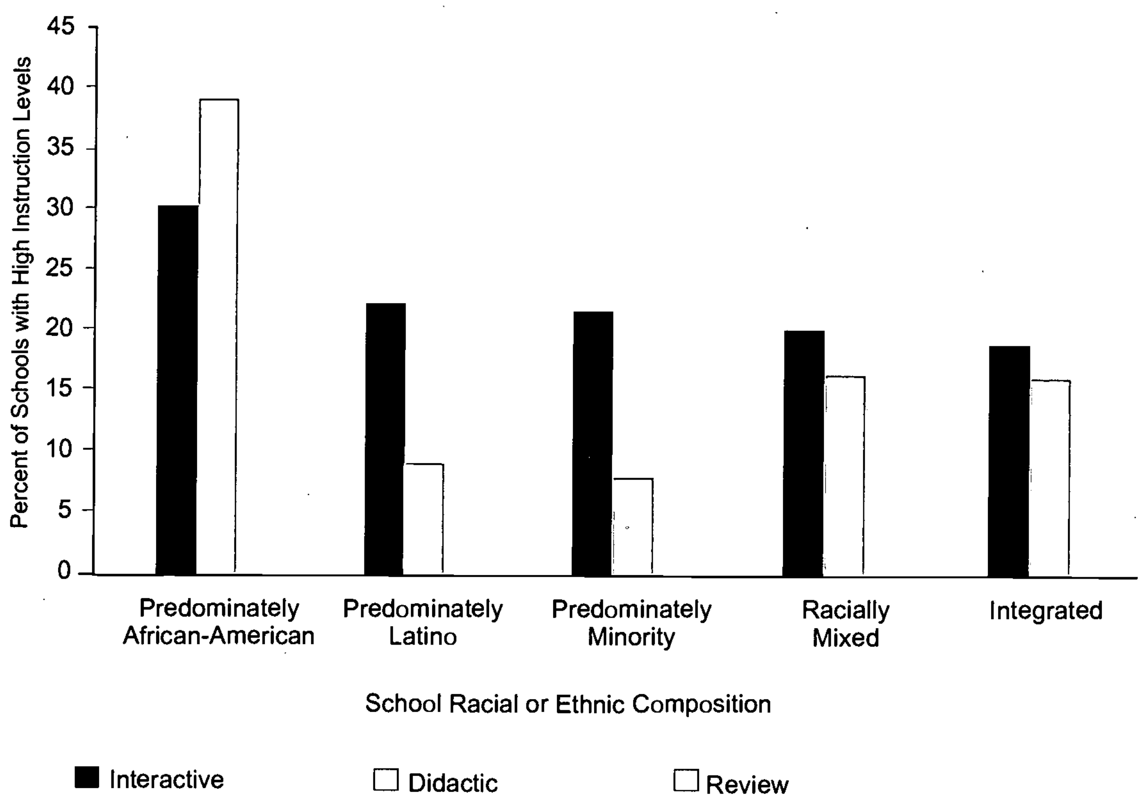
Chicago elementary students' instruction depends on characteristics of both their classrooms and their schools. Teachers use

interactive instruction more often in lower grades, but increase their use of didactic and review approaches in higher grades. Interactive instruction is more common, and didactic and review less common, in classrooms with a more favorable student composition: where students' achievement is higher, where behavior problems are fewer, and where attendance is regular.

In general, the associations between teachers' use of the three instructional approaches and school contexts are similar to our classroom findings. Teachers used interactive instruction more often in smaller schools, with fewer low-income students, and higher average achievement. Review was more common in the largest schools, schools with higher proportions of low-income students, and in schools with the lowest average achievement.

Figure 8

Didactic Instruction is Especially Common in Predominately African-American Schools



III. The Effectiveness Question: What Impact Do Different Instructional Approaches Have on Student Learning?

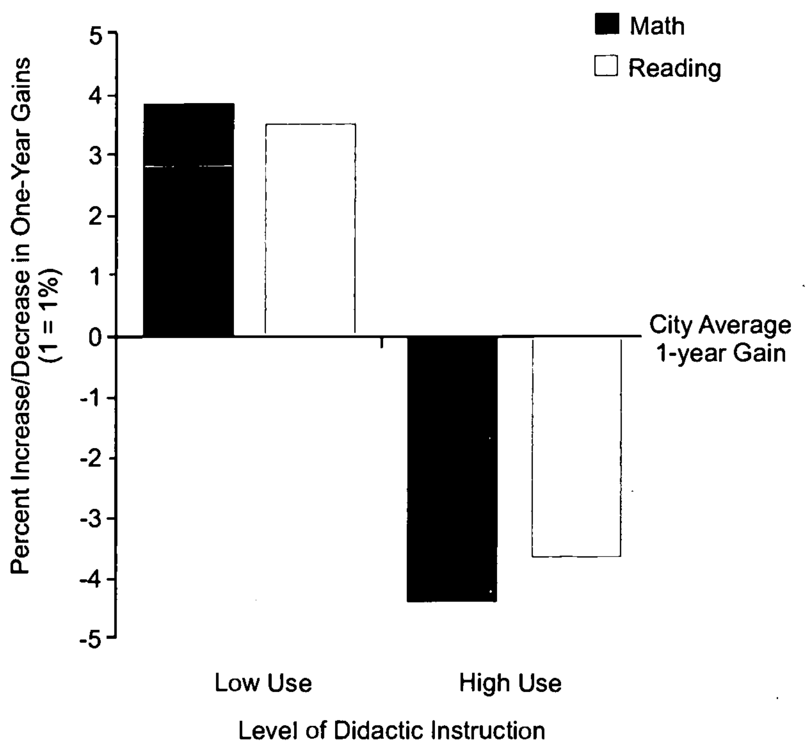
From the results presented so far, it is clear that where (and with whom) a child goes to school can make a big difference in the kind of instruction he or she receives in Chicago. Taking these differences into account, we now consider whether different approaches to instruction translate into differences in how much a student learns over the course of a year.

As discussed in the introduction, there are conflicting views about how instruction might influence student performance, particularly on standardized tests of basic skills. One view argues that interactive instruction takes students' time and energy away from the drill and repetition needed to maximize performance on these assessments. The other maintains that students need to work with information in a variety of contexts and challenges for it to become truly integrated (and thus usable) knowledge. In addition, we find conflicting views about the use of review in instruction. On one hand, it can be argued that it makes little sense to introduce new material if students have not mastered prior skills. On the other hand, because only a limited amount of time is available for instruction, the time teachers spend reviewing old material limits the time they can devote to new material. More review means less instruction on content on which students may be tested.

We examined the relationship between instruction and learning measured by student performance in reading and mathematics on the Iowa Test of Basic Skills. We defined learning as the gains each student made in his or her scores in each subject between spring 1996 and spring 1997. Our analyses take into account several student characteristics

Figure 9

Students Learn More When Use of Didactic Instruction is Low: One-Year Gains in Math and Reading Achievement, Compared to Chicago's Average



(race, gender, academic history of retention), characteristics of their classrooms (grade level, average disruption, average ability, average problem behavior), and characteristics of their schools (size, ethnic composition, percent low-income, and school average achievement).[8] The estimated effects for each instructional approach reported below also take into account how frequently the other two approaches were used.

The Influence of Didactic Instruction on Learning

In general, frequent use of didactic instruction is associated with less student learning in both reading and mathematics (see Figure 9). In schools where teachers used didactic instruction often, students learned 3.9 percent less in math and 3.4 percent less in reading than the city average. Conversely, students learned 4.4 percent more in math and 3.7 percent more in reading in schools where this type of instruction was used less often.

Although these numbers may appear small, these results represent how exposure to didactic instruction affects achievement gains in these two important subjects over only a single year. The cumulative impact of being in classrooms where teachers use didactic instruction frequently over several years, compared to classrooms where teachers use didactic instruction rarely, is substantial.

Over the eight years of elementary schools, the total impact could amount to more than a half-year less learning in schools where didactic instruction is heavily employed. In sum, the overall trend is clear: in the average Chicago elementary school, the more frequently didactic instruction is used by teachers, the less students learn in mathematics and reading over the course of a school year.

The Influence of Interactive Instruction on Learning

In contrast, consider how interactive instruction influences gains in achievement (see Figure 10). Results here show that the more frequently teachers used interactive instruction, the more their students learned in both reading and mathematics. Not only was the influence of interactive instruction on learning in the opposite direction from that of didactic instruction, this influence was consistently positive in both subjects.

In schools where teachers used interactive instruction frequently, students learned 5.1 percent more than the city average in mathematics and 5.2 percent more in reading. However, in schools where interactive methods were used less frequently, students learned 4.5 percent less both in mathematics and in reading. Again, although the one-year advantage may seem rather small, the effect can accumulate over

Figure 10

Students Learn More When Use of Interactive Instruction is High: One-Year Gains in Math and Reading Achievement, Compared to Chicago's Average

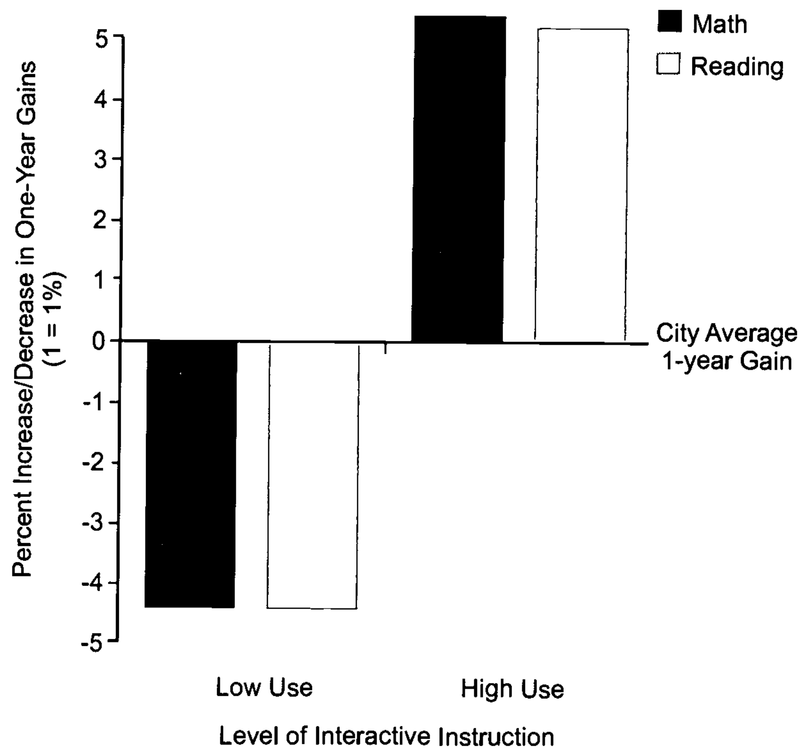
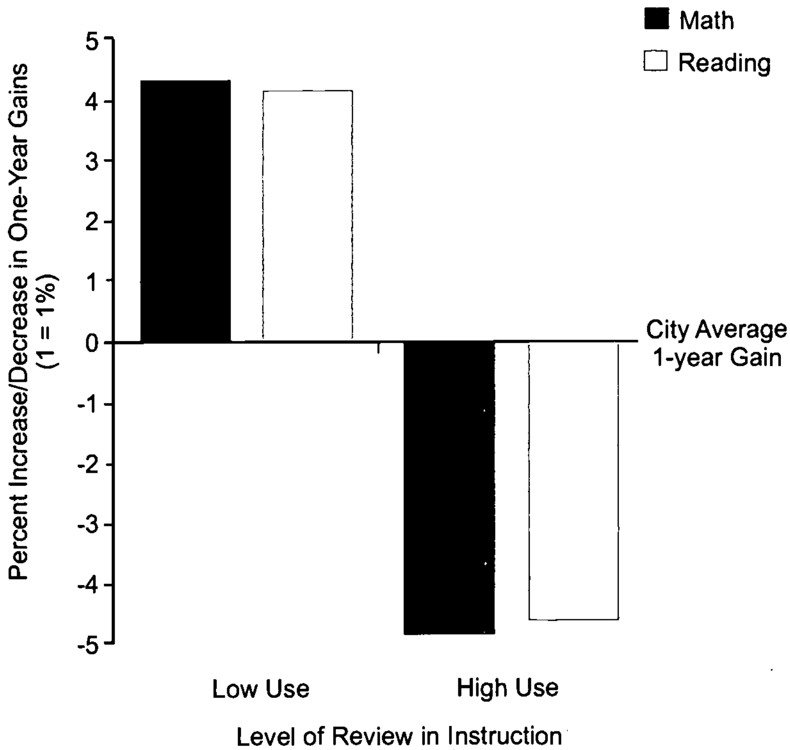


Figure 11

Students Learn More When Instructional Review is Limited (One-Year Gains in Math and Reading Achievement, Compared to Chicago's Average)



many grades. Cumulatively over the eight elementary school grades, the effect in mathematics amounts to more than half of an additional year of learning.[9]

The Influence of Review on Learning

The results for students' exposure to review are similar to findings for didactic instruction. As shown in Figure 11, the more teachers reported using review, the less their students learned in a year. Moreover, the influence of review on achievement gains was relatively similar in both reading and mathematics.

Students exposed to less frequent use of review learned 4.2 percent more in mathematics and 4.1 percent more in reading than the city's average achievement gain. Conversely, students exposed to significantly more review learned 4.8 percent less in reading and 4.9 percent less in mathematics. Although reviewing familiar content may help build a solid knowledge base for new learning, this could also diminish learning—by taking time away from teaching new material. The results support the latter interpretation, because students learned less as the frequency of review increased.[10]

IV. How Can Schools Promote More Interactive Instruction?

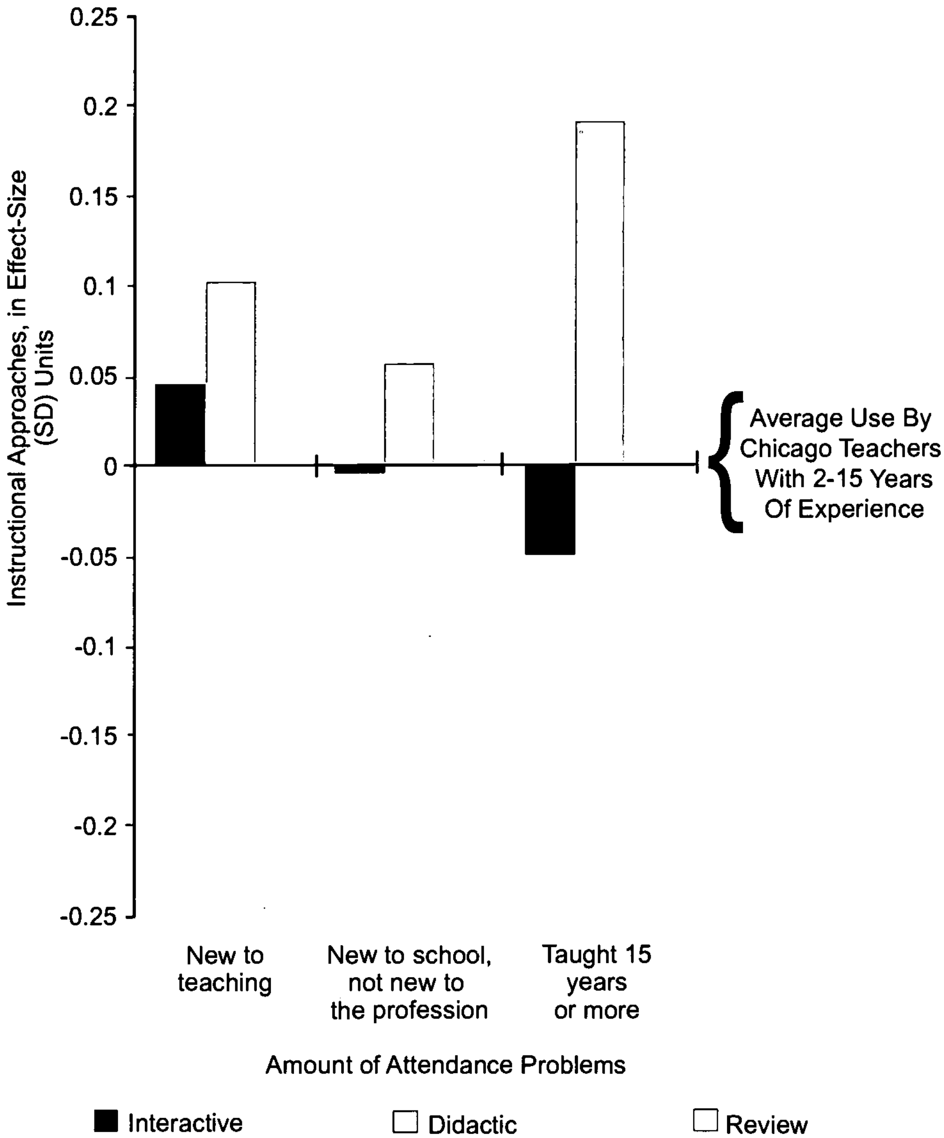
Having established that exposure to alternative approaches to instruction influences how much students learn, we decided to look more closely at how aspects of school organization and policy contribute to instructional productivity. Identifying supportive characteristics associated with more effective instruction can suggest leverage points for change, with the aim of providing all students equal opportunities to benefit from effective approaches to instruction. For example, if teachers with more advanced training were more likely to emphasize interactive instruction, increasing teacher training would be indicated as a reform strategy.

Human Resource Supports

Many characteristics of teachers might influence their use of the three instructional strategies. Here we focus on two key policy levers: teaching experience and formal professional preparation. The analyses that explore the connection between these teacher characteristics and the instructional approaches they use took into account teachers' gender, their race/ethnicity, as well as the elements of classrooms and schools described earlier in this report. For upper-grade teachers, the analyses also took into account whether the teacher was a member of an academic department. The results, displayed in Figures 12 and 13, are presented in effect-size units.

Teaching experience. Teachers were initially divided into three groups: those in their first year of teaching, those who had taught between two and 15 years, and those with more than 15 years of experience.[11] Our preliminary analyses, however, identified a distinct fourth subgroup—teachers with more than one year of experience, but who were new to their current school. Figure 12 reveals substantial differences among these four groups of teachers in their use of the three instructional approaches.

Figure 12
Association Between Teacher Experience and Use of Different Approaches to Instruction



of experience (right-hand bars in Figure 12). Of most concern are the responses from teachers new to a school, but not new to the profession. These teachers reported using interactive instructional approaches least often, and didactic methods and review most often (middle set of bars). Moreover, it is important to note that these latter results are "net of" the teaching experience effects also represented in Figure 12. That is, the reported instructional methods by a teacher new to a school and with more than 15 years of experience would be the sum of the result for the middle and right-hand bars in Figure 12. Our analyses indicate that such teachers make very little use of interactive methods, and rely heavily on didactic instruction and review.

New teachers use interactive instruction and review about as often as their faculty counterparts with 2-15 years of experience; but they use didactic strategies somewhat more frequently (left-hand set of bars in Figure 12). In contrast, very experienced teachers (with more than 15 years of experience) used interactive approaches less often than teachers with 2-15 years

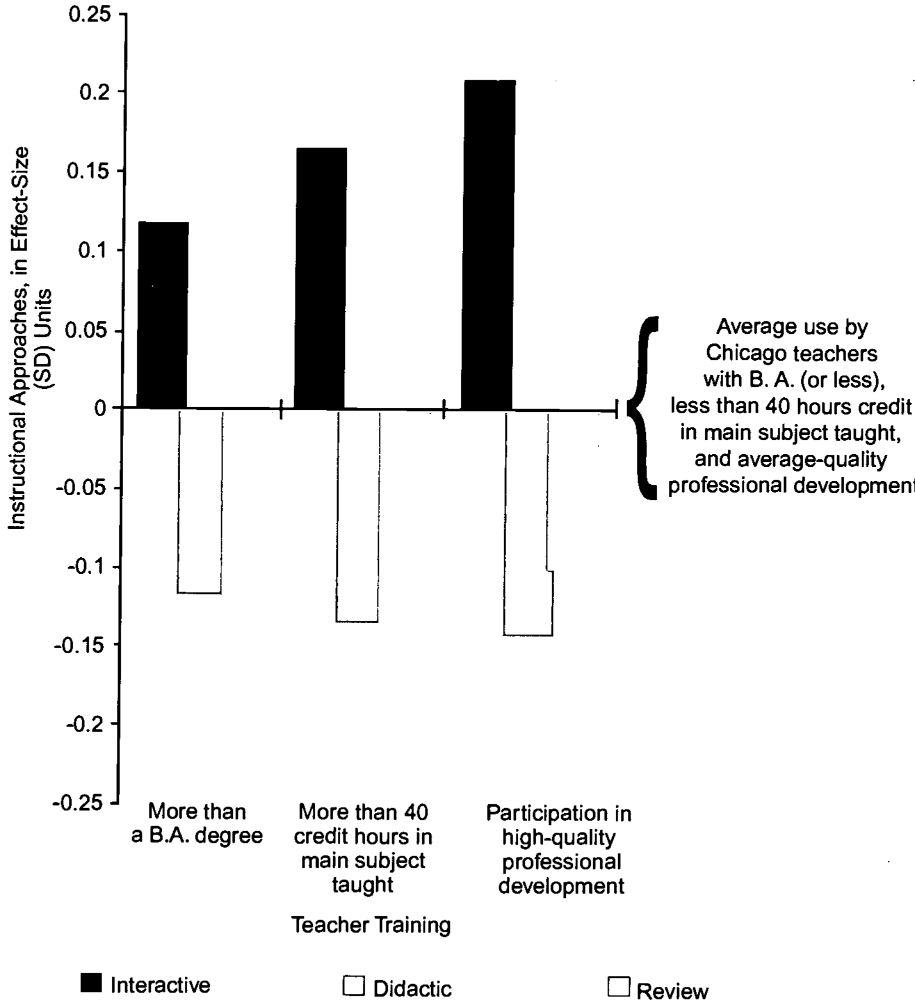
In sum, teachers' experience was associated with their usage of different instructional approaches. Individuals new to teaching, as well as other faculty colleagues with 2-15 years of experience, were most likely to engage in interactive instruction. In contrast, teachers new to their school (though not new to teaching) and very experienced teachers used interactive

methods least frequently. Since very experienced teachers comprise 45 percent of the teaching staff in Chicago schools, and those new to their schools but not new to teaching represent another six percent, these results are troubling. Over half of the teachers in the Chicago work force are not especially oriented toward more effective teaching methods. Moreover, these conclusions link with findings about professional preparation and instructional use, discussed next.

Professional preparation. To investigate whether teachers' use of instructional methods was related to their preparation as teachers, we considered three different types of preparation. First, teachers who had earned a degree beyond a bachelor's (such as a master's, specialist, or doctoral degree) were compared to those who had only a bachelor's degree or less. The second comparison contrasted teachers who had accumulated more than 40 college credit hours of instruction in their subject specialty with those who had had less training in their subject.[12] The third element of teacher preparation considered was teachers' reports of their involvement with quality professional development activities. Figure 13 displays the relationships between these measures of teacher preparation and how often they reported using the three approaches to instruction.

Figure 13

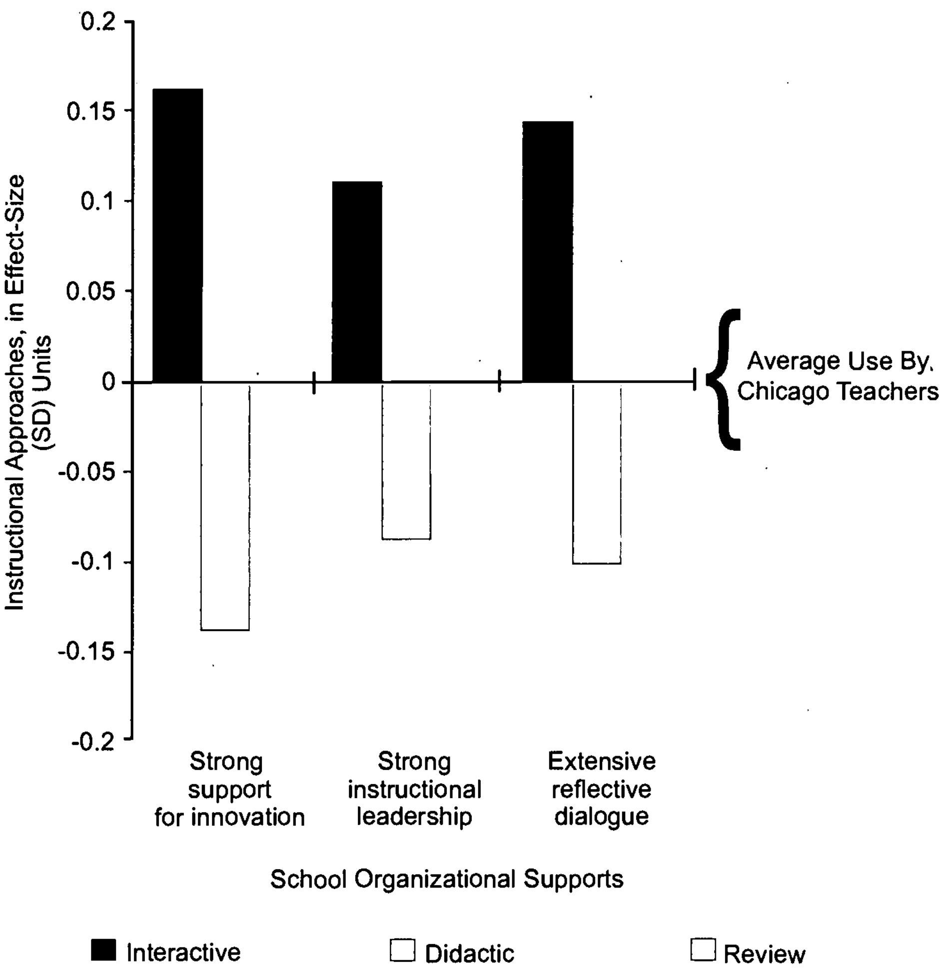
Association Between Teacher Training and Use of Different Approaches to Instruction



These three types of training were associated with instructional usage.[13] Teachers with advanced degrees used interactive methods more often and didactic methods less often, and they also reported spending much less time in reviewing old material than teachers without a degree above a bachelor's (left-hand bars in Figure 13). Similarly, teachers who completed more than 40 college credit hours in their subject area used interactive methods more often, and didactic instruction and review less often, than teachers with less coursework in their subject (middle bars). Of the

Figure 14

Association Between School Organization Supports and Use of Different Approaches to Instruction



learn. These findings support policy efforts to educate teachers to use interactive methods more frequently, either through their university pre-service education or through high-quality professional development after they enter the teaching profession.

Key Organizational Supports

Finally, we considered how aspects of school organization measured in the 1997 Consortium surveys might influence instructional usage. Three key findings emerged: More effective instruction occurs in schools where teachers are oriented toward innovation, where teachers engage in reflective discussions about their practice,

three types of preparation, teachers' reports of the quality of their professional development were most strongly associated with their use of interactive instruction (.21 effect-size units).

There is a clear trend here: Chicago elementary school teachers with more professional preparation used interactive instruction more often, and didactic instruction and review less often. The results indicate that students with better-prepared teachers had greater exposure to interactive teaching. Because interactive instruction is associated with more learning, we conclude that teachers with better professional preparation provide students with more opportunities to

and where principals demonstrate leadership in instruction. Analyses structured like those exploring teacher characteristics (Figures 12 and 13) compared schools that were high versus low on each organizational support.[14] Results in Figure 14 are again presented in effect-size units.[15]

In schools oriented toward innovation, teachers are continually learning about new ideas, the school staff has a "can do" attitude, and teachers are encouraged to change. The use of interactive instruction is more common in such places and the use of didactic instruction and review less so.

It was no surprise that we found a similar association between teachers' perceptions of their principal as an instructional leader and teachers' instructional approaches. In schools where teachers feel that the principal demonstrates strong instructional leadership, the use of interactive instruction is more common and didactic instruction and review less so. Similarly, in schools where teachers frequently engage in conversations with their colleagues about their practice, interactive instruction is used more often and didactic methods and review less often.[16]

Assessing School Performance With Different Measures

In another report from the Chicago Annenberg Research Project, "The Quality of Intellectual Work in Chicago Schools: A Baseline Report," the authors describe some indicators of high-quality intellectual performance for Chicago's elementary school students (Newmann, Lopez, and Bryk 1998). That report focused on a more direct method of evaluating the quality of the intellectual activities Chicago students engage in by examining the actual work that students undertake in their classrooms. Thus, the assessments described in that report are closely tied to the instruction that occurs in those settings. That study focused on assignments in mathematics and writing in third, sixth, and eighth grades, examining tasks that teachers rated as both "typical" and "challenging." Researchers evaluated the quality of these assignments—and the work students produced from these assignments—based on what could or should be expected for students at those grade levels. Three general criteria were used in that study to evaluate the assigned tasks against specific standards for high-quality (or "authentic") intellectual work: construction of knowledge, disciplined inquiry (that builds on a knowledge base and requires in-depth understanding and elaborated communication), and value beyond school.

The reported results were both discouraging and encouraging. Discouraging was the finding that the work Chicago students are typically assigned, and students' performance in response to those assignments, is generally of low intellectual quality. More encouraging was the finding that what teachers considered to be challenging tasks are of higher intellectual quality. Moreover, the work students produce in response to such challenging tasks was evaluated as more authentic. The conclusion drawn by that report's authors was that Chicago students can produce work of high intellectual quality when challenged to do so and when classroom instruction is organized to support it.

The measures of student performance used in this report, students' annual scores on the ITBS, do not assess performance based on the criteria laid out by Newmann and his colleagues. Rather, multiple-choice standardized tests such as the ITBS call for memorized information, retrieval of given information, or application of routine computational procedures. In one sense, it would seem that the ability to assess the effects of different instructional approaches on student learning—the major aim of this study—is compromised by focusing on assessment tools to measure learning that are more sensitive to the effects of didactic than interactive instruction. On the other hand, these are the assessments used by the Chicago Public Schools to evaluate student learning. Because students' scores on the ITBS tests are used to mandate retention, summer school attendance, and school probationary status, they are high-stakes tests. We argue that students' scores on these tests, and their progress on the tests over time, represent a realistic (albeit less authentic) way to evaluate how instruction influences learning, given the importance of such tests in the Chicago context.

V. Interpretive Summary

The aim of instruction is to promote learning. This report was motivated by a concern about how different approaches to teaching affect students' acquisition of basic skills. We found clear and consistent evidence in Chicago's elementary schools that the type of instruction teachers use influences how much students learn. In both reading and mathematics, interactive instruction had positive effects on one-year gains in achievement (from 1996 to 1997) compared to the Chicago average, whereas both didactic instruction and review had negative effects.

On average, high levels of interactive instruction added five percent to the average one-year gain in achievement for Chicago students in reading and mathematics. In contrast, high levels of didactic instruction diminished Chicago students' average in both reading and mathematics by four percent; extensive review also diminished students' gains in both subjects by another four percent. Moreover, these instructional effects work in tandem in the classroom. For example, consider the implications of our results for two contrasting prototypes of teaching—strong interactive practice with modest use of didactic and review, versus heavy didactic instruction with extensive review and limited interactive use—our findings convert into a 20 percent or more difference in learning per year between the two contexts. Indeed, type of instruction matters!

The positive effects of interactive teaching should allay fears that it is detrimental to student achievement of basic skills in reading and mathematics. Conversely, the findings call into serious question the assumption that low-achieving, economically disadvantaged students are best served by emphasizing didactic methods and review. Our results suggest precisely the opposite: to elevate mastery of basic skills, interactive instruction should be increased and the use of didactic instruction and review moderated.

On balance, our findings do not imply that didactic teaching and review should be abandoned in favor of an exclusive reliance on interactive methods. Because the analyses here relied on survey information about instruction, where teachers reported their typical use of a wide array of practices, the study is unable to describe precisely how (or how often) teachers may combine interactive, didactic, and review methods to teach specific types of content. Such inferences could be drawn only by closely scrutinizing particular teachers, and the students in their classes, over an extended period of time, as well as asking teachers to describe their rationales for using particular instructional approaches in particular instances. Clearly, some educational objectives are better accomplished through didactic and review methods. In general, effective classroom practice balances the three approaches, depending upon instructional objectives and the classroom context.

In our view, the most prudent use of these findings would be to support efforts to increase interactive teaching, aiming for a more effective balance than is currently the case. The effectiveness of interactive methods is supported by substantiated theory on how students learn. Although conventional drill, practice, and recitation can be useful, interactive methods strengthen learning because they engage students in deeper and broader thinking about subject matter. More intensive thinking helps students internalize their lessons, which increases retention. Furthermore, by encouraging students to think and reason for themselves, interactive processes build students' skills to apply old knowledge to new situations, especially to test questions that may differ slightly from those studied in class.

Although we have focused here on evaluating the effects of interactive teaching on students' basic skills achievement, we suspect that interactive teaching would have even more powerful effects on assessments that require more complex intellectual work than that measured by the Iowa Tests of Basic Skills. For example, assessments that demand more higher-order thinking, more conceptual understanding, and elaborated forms of communication would probably be even more sensitive to the level of interactive teaching that students experience.

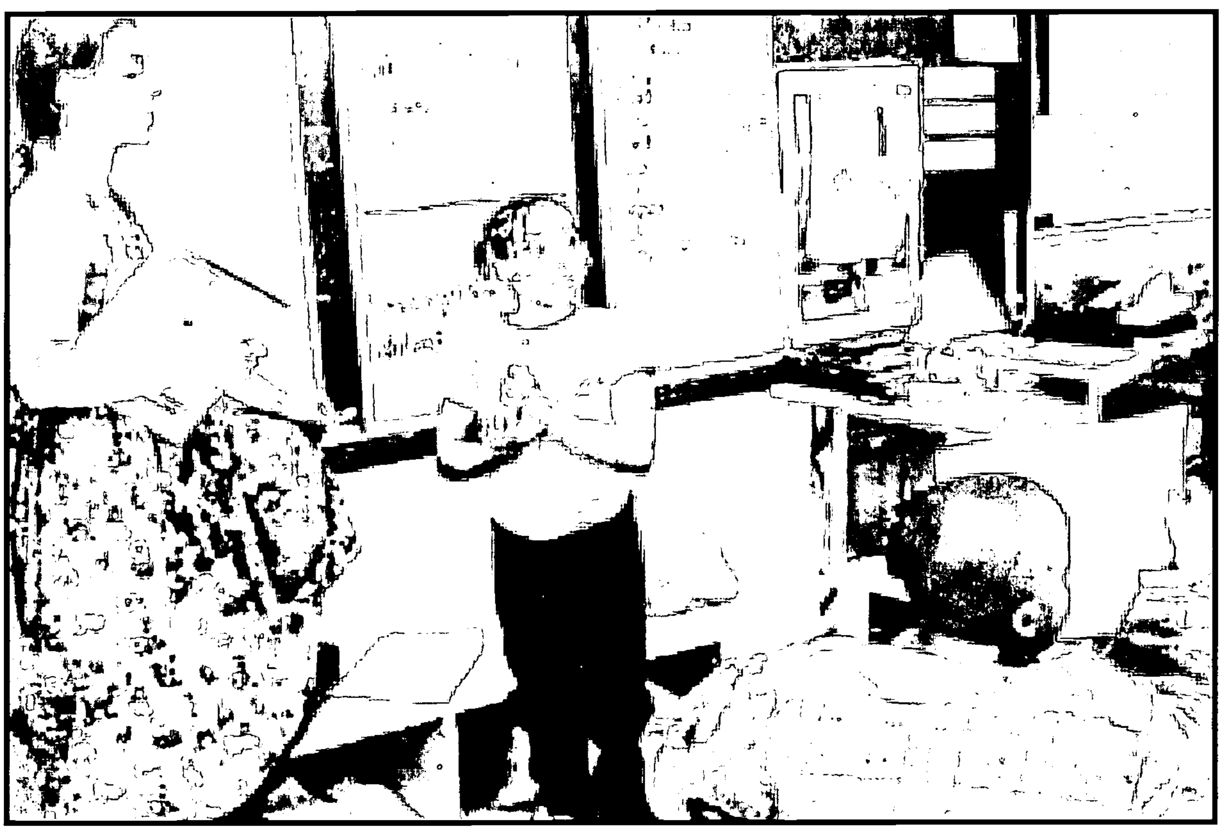
Our findings about the teachers, classrooms, and schools that are most and least likely to support interactive instruction also have important implications for how Chicago might promote more effective teaching in its elementary schools. Professional preparation and intense professional development provide one key. Our results suggest that teachers with more preparation more often use teaching practices that we have shown advance their students' learning. The persistent problems of low achievement in Chicago imply that many teachers could benefit from more training in interactive methods. Such teaching requires complex and subtle forms of engagement with students, and mastering the methods requires considerable teacher practice and support. As other researchers who study instruction have observed,

> Teaching for understanding assumes substantial new learning on teachers' part; it requires change not only in what is taught but in how it is taught. Learning how to involve students actively in the construction of knowledge, how to move beyond fact-based concepts of knowledge and learning outcomes, and how to fashion new classroom roles and relationships involves more than simply sharpening up teaching skills or teachers' professional knowledge base as conventionally conceived. Teaching for understanding requires teachers to have comprehensive and in-depth knowledge of subject matter, competence in representation and manipulation of this knowledge in instructional activities, and skill in managing classroom processes in a way that enables active student learning (McLaughlin and Talbert, 1993, p. 2).

Our findings also suggest how schools can be better organized to support interactive instruction. We found such teaching more often in schools with strong instructional leadership and support for innovation, where ample opportunities exist for teachers to observe and learn from each other about improving their practice. Although reducing school size is neither easy nor straightforward, our findings (coupled with findings that favor smaller schools in many reports from the Consortium for Chicago School Research) sug-

gest that smaller elementary schools, with less than 350 students, can better support teacher efforts to engage in interactive instruction.

In closing, we stress that this research indicates that student achievement could improve further in Chicago schools if teachers would use interactive teaching methods more often. We recognize that such an undertaking is easily urged but implemented only with great difficulty. Our most troubling findings are that students who are most in need, and who might benefit the most from being exposed to interactive instruction—academically weak students in classrooms with problematic contexts—are the least likely to be exposed to it. Redressing this mismatch is a key to moving forward on Chicago's goals: to leave no one behind and to provide more challenging work for all students.



John Booz

Appendix

37

Further Details on the Methods Used in the Analyses Reported Here

Sample Description

Students. The sample consisted of 110,775 Chicago elementary school students with ITBS test scores in mathematics and reading/language arts in both 1996 and 1997. In order to estimate achievement gains, we restricted our sample to students in grades two through eight. On average, there were 47 students per grade per school, but the range was wide (from one to 295). Of these students:

- 52 percent were female;
- 64 percent were African-American, 23 percent Latino, 3 percent Asian, 11 percent white;
- 10 percent had repeated a grade at least once.

Teachers. All full-time classroom teachers who responded to the 1997 teacher survey were included in the analysis. The sample varied according to which instruction measure was being considered, as follows:

- Interactive instruction: 3,284 teachers in language arts, 2,258 in mathematics, 5,542 for the combined measure.
- Didactic instruction: 3,328 teachers in language arts, 2,258 in mathematics, 5,586 for the combined measure.
- Review: 5,586 teachers included in this composite.

Typically, there were about four teachers per grade per school, with responses evenly divided between the language arts and mathematics portions of the survey. Of these teachers:

- 15 percent were male;
- 55 percent were either African-American or Hispanic;
- 47 percent had earned degrees past a bachelor's;
- 34 percent had taken more than 40 hours of post-secondary training in their major subject area;
- 4 percent were new to teaching, and 45 percent had been teaching for more than 15 years;
- 22 percent were affiliated with a department.

Schools. Our school sample consisted of 384 schools with teacher survey information available for the analysis of mathematics gain; 380 schools were included in the analysis of reading gain. Of the 380 schools:

- The average enrollment was 664 students, with 11 percent of the schools smaller than 350 and 37 percent larger than 700 students;

- The average mobility rate was 28 percent, the average truancy rate was two percent, and the average daily non-attendance rate was seven percent. We combined these variables to measure the level of student instability for the school.
- 47 percent of the schools were predominantly African-American, 9 percent predominantly Hispanic, 15 percent predominantly ethnic minority, 12 percent racially mixed, and 17 percent integrated.

Measures Used in Analysis

Achievement. We used the ITBS mathematics and reading tests for 1996 and 1997. These tests were converted to Rasch-scaled estimates, to provide a continuum on which to examine change. For our analyses, we examined the change or gain in test scores from 1996 to 1997 in the two subjects. In order to report our findings in terms of a year's gain, we first subtracted the 1997 from the 1996 ITBS Rasch-adjusted test scores.

We then ran a three-level HLM analysis with these gains in tested performance as the outcomes, estimated for students nested within grades nested within schools. This analysis showed that, across all Chicago students in grades two through eight, the average Rasch estimated gain in mathematics achievement from 1996 to 1997 was .67; the average gain in reading was .70. Our results are reported in the metric of "one year's gain," which we computed by dividing the estimated instructional effects in the mathematics analyses by .67 and the estimated instructional effects in reading by .70. These results represent the proportion of the average gain between 1996 and 1997 that can be attributed solely to each instructional approach—didactic instruction, interactive instruction, review.

Instruction. The three measures of instruction were constructed, using Rasch scaling methods, from teacher responses to items in the 1997 Consortium survey. The individual items included in these scales are listed in Appendix Table A-1. Our analyses used these three composite measures of instruction in two ways: (1) as grade-level aggregates to examine their effects on student achievement gains, and (2) as teacher-level outcomes in the analyses that investigated (a) students' exposure to different forms of instruction, described in our classroom and school-context analyses, and (b) our analyses examining human resource supports for instruction.

Based on both existing research and our knowledge about didactic and interactive instruction, we divided survey items, drawn from questionnaires completed by teachers in grades two through eight, into scales. The items came from three sections of the 1997 teacher survey:

- Those to which all teachers responded;
- Those completed by teachers about their instruction in mathematics;
- Those completed by teachers about their instruction in reading or language arts.

Upper-grade teachers with specialized teaching assignments answered questions based on that specialty. Lower-grade teachers, most of whom did not specialize, were divided randomly into two groups—one responded to mathematics items and the other to language arts items. Thus, all teachers completed two of the three item sets. In order to align subject-matter instruction with achievement in the same subject, we eliminated from our sample teachers who reported teaching neither reading nor math. When these measures were used as outcomes in our HLM analyses, the interactive and didactic scales for mathematics and reading were each combined into

a single interactive and a single didactic scale. Scale reliabilities for the instructional measures, measured on teachers, are as follows: interactive (reading): .82; interactive (math): .87; didactic (reading): .82; didactic (math): .71; and review: .68.

To estimate the effect of instruction on gains in achievement, we created an aggregate of the teachers at a given grade level for each school. Ideally, we would have liked to link each student to his or her teacher in the 1996-97 school year, but this information was not available in the survey data. However, we are able to locate each student and teacher within a grade at a particular school. Therefore, the closest we could get to the actual instruction received by each student is the average delivered in his or her school and grade level.

More specifically, our three-level HLM models included a measurement model on teachers nested within grade levels. As the instruction scales were constructed using Rasch scaling, the estimate of each teacher's instructional approach also included an individual standard error of measurement associated with the estimate. These standard errors were taken into account in Level 1 of the HLM analysis. In addition, we introduced controls for teachers' gender and race or ethnicity at Level 2, because preliminary analyses suggested a systematic bias in teacher reports of instructional usage associated with these two characteristics. This HLM analysis produces an empirical Bayes estimate at Level 3 for each grade's average instructional usage for each type of instruction in each school. We used this process separately for the language arts and mathematics interactive and didactic measures, using content-specific instructional measures for the reading and mathematics achievement outcomes.

Classroom achievement level. On the survey, teachers were asked to identify the type of student in their target class (the class they were reporting on for their instruction (TGT53Q01). They were asked to describe their students as: entirely below grade level, below or at grade level, at or above grade level, or entirely above grade level. The distribution of these responses was as follows:

- 33.3% of teachers responded "below grade level";
- 43.5% of teachers responded "below or at grade level";
- 21.4% of teachers responded "at or above grade level";
- 1.9% of teachers responded "above grade level."

We began by constructing three groups: (1) teachers who identified their class as entirely below grade level; (2) teachers who identified their class as below or at grade level; and (3) those who reported their class to be either at and above or entirely above grade level. From these groups, we formed two dummy variables, which measured below-average and at or above-average classes. Used together in the analysis, both group one and group three were compared to group two. In the instruction-as-outcome analyses, these two dummy variables measured an important aspect of the classroom context. In the achievement gain outcome analyses, we used grade-level aggregates of these measures, which captured the percentage of teachers across grades in each school who reported their class as entirely below grade level or entirely above grade level.

Classroom behavior. On the survey, teachers were asked to indicate the proportion of new students and students who had left the class since the beginning of the year who exhibited the following behaviors:

- Serious reading problems (TGT58Q01 & TGT60Q01);
- Lacked knowledge or skills (TGT58Q02 & TGT60Q02);
- Created behavior problems (TGT58Q04 & TGT60Q04).

These items scaled well together, forming a composite measure that profiles classrooms with different types of "problem" behaviors. The composite was converted to a z-score (M = 0, SD = 1.0). The composite scale had a scale reliability of .87.

Irregular classroom attendance. On the survey, teachers were asked to indicate the average number of:

- Absent students (TGT56Q01);
- Tardy students (TGT56Q02);
- New students enrolled since October (TGT57Q01);
- Students who had left the class since October (TGT59Q01).

These items scaled together, forming a composite profile of the level of classroom "upheaval" with which the teacher had to contend. The composite was converted to a z-score (M = 0, SD = 1.0). The composite had a scale reliability of .65.

Quality professional development. This measure, which was constructed using Rasch scaling methods, included items drawn from teachers' survey reports about their experiences with professional development. This teacher-level measure was converted to a z-score (M = 0, SD 1.0). We defined "low" quality as one standard deviation below the mean on this measure, and "high" quality as one standard deviation above the mean. This measure, with a scale reliability of .91 at the teacher level, consisted of items where teachers described their professional development experiences as those that:

- Provided opportunities to work productively with teachers from other schools;
- Changed the way teachers talk about students in this school;
- Included opportunities to think carefully about, try, and evaluate new ideas;
- Shifted approaches to teaching in this school;
- Helped my school's staff work better together;
- Addressed the needs of the students in the classroom;
- Deepened my understanding of subject matter;
- Helped me understand my students better;
- Were sustained and coherent, rather than short-term and unrelated;
- Included opportunities to work with colleagues in my school;
- Led me to make changes in my teaching.

Instructional leadership. This school-level composite measure, constructed using Rasch scaling methods, was drawn from teachers' reports about the principal at their school that year. The composite was aggregated to the school level and converted to a z-score (M = 0, SD = 1.0). We defined as "low" schools that were one standard deviation or more below the mean on this measure, and as "high" schools which were one standard deviation or more above the mean. The measure, with a scale reliability of .90 at the teacher level, included the following items in which teachers reported whether their principal:

- Carefully tracks students' academic progress;
- Understands how children learn;
- Presses teachers to implement what they have learned in a professional manner;
- Communicates a clear vision for our school;
- Sets high standards for student learning;
- Sets high standards for teaching;
- Makes clear to the staff his or her expectations for meeting instructional goals.

Orientation toward innovation. This school-level composite measure, constructed using Rasch scaling, includes items in which teachers described the other teachers in their school. The composite was aggregated to the school level and converted to a z-score (M = 0, SD = 1.0). We used "low" to indicate schools that were one standard deviation or more below the mean on this measure, and "high" for those schools which were one standard deviation or more above the mean. The scale, with a reliability of .89 at the teacher level, included the following items:

- Proportion of teachers willing to take risks to make the school better;
- Proportion of teachers eager to try new ideas;
- Teachers in this school have a "can do" attitude;
- All teachers are encouraged to "stretch and grow";
- Teachers are continually learning and seeking new ideas;
- Proportion of teachers really trying to improve their teaching.

Reflective dialogue. This school-level composite measure, constructing using Rasch scaling, included items in which teachers described other teachers in their school. The composite was aggregated to the school level and converted to a z-score (M = 0, SD = 1.0). We used "low" to indicate schools that were one standard deviation or more below the mean on this measure, and "high" for those schools which were one standard deviation or more above the mean. The scale, with a reliability of .80 at the teacher level, included the following items:

How often teachers have had conversations with colleagues about:

- The goals of the school;
- Development of new curriculum;

- Managing classroom behavior;
- What helps students learn best.

Teachers agree that other teachers in the same school:

- Regularly discuss assumptions about teaching and learning;
- Share and discuss student work with other teachers;
- Talk about instruction in teachers' lounge, faculty meetings, etc.

Structure of HLM Models

Achievement gain analyses. The HLM analyses that investigated instructional effects on achievement gain examined these effects on student learning within grades within schools. As noted above, we could not identify the classroom each student was in, but we were able to locate students within grade levels in each school. Given the relatively high variability among teachers within schools on these measures, this limitation implies that the effects that we uncover in this study may well be underestimates of the true structural parameters.

In the HLM analyses on achievement gains, we used the difference in ITBS scores in mathematics and reading between 1996 and 1997 as the outcome. The analyses had a three-level structure. Level 1 (students) included dummy-coded controls for students' gender, minority status, poverty status, and retention history. These controls were all fixed (i.e., their between-group variance was set to zero) and grand-mean centered. Level 2 (grade level) contained the three instructional variables: didactic, interactive, and review. These measures were fixed and grand-mean centered. Also at this level, we included a series of grade-level dummy variables, for grades two through eight, with grade five as the excluded category, as well as averages across grades of teachers' reports of their classroom contexts (class ability level, classroom behavior, and classroom attendance). Again, as controls, these variables were fixed and grand-mean centered. Level 3 (school model) included controls for school size, school racial or ethnic composition, a measure of school instability (this included truancy and mobility), and the percent of low-income students at the school. Thus, effects of each form of instruction were estimated while taking teachers' scores on the other two instruction variables into account, as well as characteristics of students, teachers, classrooms, and schools. The critical results from this model reported in the study were the relationships between each instructional measure and the gain in achievement.

Instruction-as-outcome analyses. We also used 3-level HLM analyses to investigate teacher, classroom, and school effects on the frequency of instructional usage. For these analyses, the dependent measures were the individual teacher's reports of their instructional use of didactic, interactive, and review methods. As was the case in constructing the grade-level aggregates, we again controlled for the standard errors of the Rasch estimates of these outcomes (level 1). Level 2 (teacher-classroom model) included controls for teachers' gender, ethnicity, education, experience in teaching, professional preparation, experience with quality professional development, grade level, relative classroom ability, level of problem behavior, and level of student absenteeism. All level-2 variables were fixed and grand-mean centered (i.e., they were treated as control variables in estimating the other level-2 effects). Level 3 (school) included school size, school instability, school poverty concentration, instructional leadership, support for innovative practice, and whether teachers were departmentalized. Each of these variables was grand-mean centered. Due to problems of collinearity, effects of the latter two measures of key school supports were each entered separately, rather than in the presence of the other two.

Table A1

List of Items Used in Instruction Scales

Items Used to Make Didactic Instruction Scale	Hours of class preparing for student testing How often do you have students memorize How often do you lecture for more than half the period How often do you have students do workbook/textbook exercises How much use do you make of the textbook How often do you have students read silently How often do you have students read aloud How often do you see highly structured call/response questioning Lesson content emphasizes basic facts and procedures	Judge student learning with multiple choice or true/false Judge student learning with short-answer tests Judge student learning with standardized tests Assess progress based on use of proper format Assess progress based on neat and organized work Significant time spent on vocabulary Significant time spent on proper grammar and punctuation Significant time spent on handwriting
Items Used to Make Interactive Instruction Scale	How often do you assign one-week projects How often do you relate subject to student interest How often do you use cooperative groups How often do you have students brainstorm How often do you have students debate ideas Percent of time spent on studying topic in-depth Percent of time spent on relating topic to student experiences Percent of time spent on students producing original work Judging student learning with student discourse in class Judging student learning with student presentation of work Judging student learning with essay tests Judging student learning with open-ended problems Judging student learning with individual projects Judging student learning with group projects Assess student progress: demonstrate reasoning Assess student progress: original creative work Percent of time spent on differentiating fact from opinion Percent of time spent on drawing inferences Percent of time spent on understanding historical context Percent of time spent on understanding author's perspective	Percent of time spent on analyzing and interpreting literature Percent of time spent on synthesizing ideas from readings Percent of time spent on applying literature to current situations Significant time spent on integrating reading and writing instruction Emphasis on writing process methods Have students discuss readings in small groups Have students write about something they have read Emphasis on reasoning and analytic skills Emphasis on communicating math ideas Emphasis on practical everyday applications How often do you have students work with measuring instruments How often have students show their work How often have students explain how they solved problems How often have students write a proof How often have students write/do math projects How often have students apply math to life How often have students show how they arrived at answer How often have students devised a solution without standard procedure How often have students consider alternative procedures How often do you teach deep understanding concept
Items Used to Make Review and Repetition Scale	Time spent: reviewing content/skills from previous lessons Weekly teaching spent reinforcing content/skills Homework: emphasis on student review and practice of material Time in class given to start homework	

HLM Results on the Effect of Instruction on Gains in Mathematics and Reading

	Outcomes	
Independent Variables	Math Gain	Reading Gain
Instruction Effects		
Didactic	-.04**	-.03*
Interactive	-.04**	.03**
Review and repetition	-.04**	-.03**
Grade Level Controls		
Grade 2	.20***	.83***
Grade 3	.27***	.33***
Grade 4	.22***	-.23***
Grade 6	-.05	.06
Grade 7	-.33***	.06
Grade 8	-.28***	.11**
Percent below grade level	.02	.003
Percent at/above grade level	-.03	-.02
Average problem behavior	.003	.01
Irregular class attendance	-.02	.01
Student Controls		
Poverty	-.02***	-.02
Female	.003	-.01
Black or Latino	-.01	-.0004
Ever held back	-.07***	-.02
School Controls		
Small school	.04	.08
Large school	-.02	-.02
Percent low-income students	-.02	-.02**
School instability	.04*	.03
Predominantly African-American	-.10**	-.11**
Predominantly Latino	.05	.06
Predominantly minority	-.03	.001
Racially mixed	.10*	.07
Prior achievement level (1996)	.35***	.23**

* Significant at .05 or higher
** Significant at .01 or higher
*** Significant at .001 or higher

Table A3

HLM Results on the Impact of Teacher, Classroom and School Characteristics on Level of Instructional Use

Independent Variables	Outcomes		
	Interactive Instruction	Didactic Instruction	Review
Teacher Characteristic			
African-American or Latino	.35***	.31***	.09*
Male	-.10**	-.05	-.06
New to school	-.17**	.13*	.10*
New to teaching	.17**	-.12*	-.16*
Has taught 15 years or more	-.13*	.03	.04
Has more than a B.A.	.13*	-.11***	-.22***
Has more than 40 hours college training in main subject area	.17***	-.12***	-.12**
Quality of professional development	.21***	-.13***	-.09**
Teaches in a department	-.26***	.30***	.31***
Classroom Characteristic			
Grades 1-3	.31***	-.08**	-.03
Grades 6-8	-.10**	.08*	.03
All below grade level	-.23***	.13***	.70***
At or above grade level	.19***	-.17***	-.51***
Level of problem behavior	-.03~	.17***	.34***
Amount of upheaval	.004	.12***	.09***
School Characteristic			
Small school	.14*	-.04	-.09
Large school	-.03	.11**	.21**
Percent low-income students	-.09	.14**	.15*
School instability	-.02	.11	.15*
Predominantly African-American	-.07	.14*	.01
Predominantly Latino	.02	.03	.15*
Predominantly minority	-.02	.02	.01
Racially mixed	.05	.04	.09
Prior achievement level (1996)	.05~	-.07~	-.10
Support for innovation (A)	.16**	-.13**	0.16**
Instructional leadership (A)	.13**	-.10*	-.05
Reflective dialogue (A)	.15**	-.11**	-.14***

(A) These variables were entered separately, each in its own HLM analysis on the three outcomes, because they are so highly correlated that when one uses all three, none are significant.

~ p < .10 *p < .05 **p.01 *** p < .001

Endnotes

[1] The two lessons were not actually observed in the same grade in a single specific school, but they represent composite descriptions of lessons that have been observed in the Chicago Annenberg Research Project fieldwork.

[2] Good and Brophy (2000) summarized differences in the two approaches as the transmission view and social construction view of teaching. Examples of proposals or programs that reflect more emphasis on didactic instruction include Bereiter and Englemann (1966); Hirsch (1987); Stein, Silbert, and Carnine (1997); and Strickland (1998). Examples emphasizing interactive instruction include Brooks and Brooks (1993); Driver (1995); Goodman and Goodman (1979); and Thomas (1991). The Web site www.publicagenda.org provides periodic surveys of public attitudes and summaries of arguments related to these approaches; for example, under "Framing the Debate, Perspective #2: Creating Student-Centered Schools."

[3] Delpit (1995), for example, argued that certain kinds of "progressive" pedagogy, for example in teaching the writing process, neglect important needs of many low-income urban African-American students to learn basic skills.

[4] Although Figure 1 suggests a trend for the use of review that mirrors the use of didactic instruction, the coefficients associated with this grade-level comparison were not statistically significant. In Figures 1 through 4, the instructional approaches were measured in effect-size (or standard deviation) units.

[5] Irregular classroom attendance problems were measured as a combination of teachers' responses about the number of: (a) absent students; (b) tardy students; (c) new students enrolled since October 1996; and (d) students who had left the class since October 1996. More detail on this measure is provided in the appendix.

[6] To create this measure, the three instructional measures were each aggregated to the school level. These aggregates were then divided into quartiles. "High-use" schools were those where the frequency of each instructional approach was in the highest quartile. Analyses in this section are descriptive, and thus do not take into account other factors.

[7] Here we divided schools into three groups, based on students' average scores on the ITBS tests in reading and mathematics in 1996: the lowest quartile, the middle two quartiles, and the upper quartile.

[8] Unfortunately, the structure of our analyses does not allow us to link students directly to their teachers. Therefore, our statistical controls for the several measures of classroom context are averages of: (a) class ability level, (b) behavioral problems, (c) and irregular attendance across classrooms in the same grade in any school. These are included at the class level in our HLM analyses, where instructional usage is also included. The models also control for the school context factors shown in Figures 5 through 8. The details of the HLM results shown in Figures 9 through 11 are displayed in Table A-2. For each instructional practice, we defined high use as the average report among the top quartile of teachers on each respective measure. Correspondingly, low use is the average report among the bottom quartile of teachers on each measure. The effects displayed in Figures 9 through 11 compare the results for these high- and low-use groups relative to the citywide average for each instructional practice.

[9] In these analyses we also considered whether interactive instruction had a different effect on achievement gains, depending upon the amount of didactic instruction reported, above and beyond the main effects. For example, perhaps interactive instruction would be most effective when teachers also reported using a moderate amount of didactic instruction and least effective when they reported very low levels of didactic instruction. We did find a statistically significant (but very small) interaction between interactive and didactic instruction. However, even when taking into consideration this interaction, the positive effects of interactive instruction remained, regardless of the level of didactic instruction (i.e., the interactive effects were not disordinal). In short, whether didactic use was high or low, students in classes with frequent use of interactive instruction had higher achievement gains. Thus, we report only the main effects here for purposes of simplicity.

[10] We suspect that these findings underestimate the actual effects of instruction on achievement—for two reasons. First, if teachers actually engaged in less interactive instruction and more didactic than they reported (which might be the case since interactive methods are seen as more socially desirable to report), the reliability of the measures of instruction used in these analyses could be reduced. There is wide agreement among researchers that less reliable measures systematically underestimate actual effects compared to more reliable measures. Therefore, the effects estimated in these analyses would be larger if more reliable measures of instruction were available.

Second, a portion of the instructional effects reported here may be biased by a general statistical phenomenon called "regression to the mean." This phenomenon may have occurred here because students with lower scores on achievement measured in 1996 had a higher probability of recording a large gain than their initially higher-scoring counterparts. In contrast, students who scored higher at the first testing point (i.e., closer to the top score on the 1996 test) had less "room to gain." Since didactic instruction was more common in classes enrolling more students with lower achievement levels, and

interactive instruction was more common in classrooms with higher achievement (see Figure 7), the net effect here would be to overestimate the effects of didactic instruction and to underestimate interactive effects.

[11] In this sample, four percent of teachers were new to teaching. The teachers who had between two and 15 years of teaching but who were new to their current schools comprised six percent of the sample. Almost half (45 percent) of the teachers in this sample had been teaching for more than 15 years.

[12] For teachers who taught in the lower grades, the most common specialty area was reading. However, some teachers in the lower grades identified their specialty as mathematics or other elementary subjects (including science, music, and art).

[13] It is important to keep in mind that these relationships for the contextual analyses are cross-sectional rather than longitudinal (as is the case for the achievement analysis), which limits our certainty about causal direction. Although the quality and quantity of teachers' professional preparation may well influence how they deliver instruction (and various studies of teaching make a strong case for this), it is also possible that teachers more committed to interactive approaches could pursue training more actively.

[14] The key organization supports considered in this analysis were highly correlated with one another. Therefore, each was considered separately. However, each analysis took into account the elements of school demography and context, as well as demographic characteristics of teachers, the grade level, and classroom context.

[15] The results which generate this figure can be found in the HLM results in Table A-3. The effect sizes reported here were calculated by dividing each of the relevant coefficients in Table A-3 by their respective school-level standard deviations (.78 for interactive instruction, .62 for didactic instruction, and .67 for review).

[16] Other Consortium reports that have studied collaboration among school members have employed a composite measure of school-based professional community. However, in this report we chose to focus attention on a single component of professional community: reflective dialogue. Our rationale for this decision is that it is this component which more closely aligns with the teacher behaviors that are key to developing interactive methods.

References

Bereiter, Carl and Seigfried Englemann (1966). *Teaching Disadvantaged Children in the Preschool.* Englewood Cliffs, NJ: Prentice Hall.

Brooks, J. G. and Martin G. Brooks (1993). *In Search of Understanding: The Case for Constructivist Classrooms.* Alexandria, VA: Association for Supervision and Curriculum Development.

Carpenter, Thomas P., Elizabeth Fennema, Penelope Peterson, Chia-Ling Chiang, and Megan Loef (1989). Using knowledge of children's mathematics thinking in classroom teaching: an experimental study. *American Educational Research Journal,* 26, no. 4: 499-531.

Cobb, Paul, T. Wood, Erma Yackel, J. Nicholls, Grayson Wheatley, B. Trigatti, and M. Perlwitz (1991). Assessment of a problem-centered second-grade mathematics project. *Journal for Research in Mathematics Education,* 22, no. 1: 3-29.

Delpit, Lisa (1995). *Other People's Children: Cultural Conflict in the Classroom.* New York: New Press.

Driver, Rosalind (1995). Constructivist approaches to science teaching. In *Constructivism in Education.* L. P. Steffe and J. Gale (Eds.), 385-400. Hillsdale, NJ: Erlbaum.

Good, Thomas L. and Jere Brophy (2000). *Looking in Classrooms* 8th ed. Reading, MA: Longman.

Goodman, Kenneth S. and Yetta M. Goodman (1979). Learning to read is natural. In L.B. Resnick and P. A. Weaver (Eds.), *Theory and Practice in Early Reading,* Vol. 1. Hillsdale, NJ: Erlbaum.

Hirsch, Eric D., Jr. (1987). *Cultural Literacy: What Every American Needs to Know.* Boston: Houghton Mifflin.

Johnson, J. (1995). *Assignment Incomplete: The Unfinished Business of Education Reform.* A Report from Public Agenda. New York: Public Agenda.

Johnson, Jean and John Immerwahr (1994). *First Things First: What Americans Expect From the Public Schools.* A Report from Public Aganda. New York: Public Agenda.

Knapp, Michael S., Patrick M. Shields, and Brenda J. Turnbull (1992). *Academic Challenge for the Children of Poverty: Summary Report.* Washington, D. C.: Office of Policy and Planning, U. S. Department of Education.

Lee, Valerie E., Julia B. Smith, and Robert G. Croninger (1997). How high school organization influences the equitable distribution of learning in mathematics and science. *Sociology of Education,* 70: 128-50.

Levin, Martin, Fred M. Newmann, and D. W. Oliver (1969). *A Law and Social Science Curriculum Based on the Analysis of Public Issues: A Final Report.* Washington, D. C.: Office of Education, Bureau of Research, U. S. Department of Health, Education, and Welfare.

McLaughlin, Milbrey W. and Joan Talbert (1993). Introduction: new visions of teaching. In D. K. Cohen, M. W. McLaughlin, and J. E. Talbert (Eds.), *Teaching for Understanding: Challenges for Policy and Practice.* San Francisco: Jossey-Bass.

Rosenthal, Robert and Ralph L. Rosnow (1984). *Essentials of Behavioral Research: Methods and Data Analysis.* New York: McGraw Hill.

Silver, Edward and Suzanne Lane (1995). Can instructional reform in urban middle schools help students narrow the mathematics performance gap? *Research in Middle Level Eduation,* 18 no. 2: 49-70.

Stein, M., Jerry Silbert, and Douglas Carnine (1997). *Designing Effective Mathematics Instruction: A Direct Instruction Approach.* Upper Saddle River, NJ: Prentice Hall.

Strickland, Dorothy S. (1998). *Teaching Phonics Today: A Primer for Educators.* Newark, DE: International Reading Association.

Tharp, Roland G. (1982). The effective instruction of comprehension: results and description of the kamehaneha early education program. *Reading Research Quarterly,* 17, no. 4: 503-27.

Thomas, D. A. (1991). *Children, Teachers, and Mathematics.* Boston: Allyn and Bacon.

About the Authors

Julia B. Smith is an Associate Professor of Education at Oakland University. She received her Ed.D. in Education Studies from the University of Michigan. Her research focuses on untangling different sources of risk for students who have been traditionally unsuccessful in public schooling. She has recently completed a study exploring the impact of parenting behavior on child development. In addition, she has recently completed a series of studies investigating the role of instruction, curriculum content, and task complexity in predicting high school students' history test scores using the National Assessment of Educational Progress. Currently, she is studying the effects of iron deficiency on early child cognitive development, as well as the impact of behavioral comorbidity on school performance in mathematics and reading.

Valerie E. Lee is a Professor of Education at the University of Michigan and a member of the Consortium's evaluation team for the Chicago Annenberg Challenge. She received her Ed.D. in Administration, Planning, and Social Policy from Harvard University. Her research, which is located within the Sociology of Education, focuses on what schools can do, in terms of structure and policy, to increase both equity and excellence. She has recently completed a two-year qualitative study of high schools divided into schools-within-schools. Currently, she is investigating successful educational contexts for young children living in poverty, with both quantitative and qualitative methods. She recently completed two studies: (1) an evaluation of the relative efficacy of full-day and half-day kindergartens on children's learning; and (2) how personalism influences students' propensity to drop out of high school.

Fred M. Newmann is Professor Emeritus of Curriculum and Instruction at the University of Wisconsin-Madison and a member of the Consortium's lead team for the Chicago Annenberg Research Project. He has worked as a high school social studies teacher, curriculum developer, teacher-educator and researcher of school reform. Having directed the National Center on Effective Secondary Schools and the Center on Organization and Restructuring of Schools, his recent publications propose criteria for authentic achievement and assessment and describe how restructuring in schools nationwide affects authentic student achievement, equity, empowerment, professional community, and accountability. He currently directs a national study of professional development to build capacity in low-income schools. Newmann co-authored the Chicago Annenberg Research Project report, *The Quality of Intellectual Work in Chicago Schools: A Baseline Report* (October 1998).

NOTICE

Reproduction Basis

X	This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.
	This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").